BEST AVAILABLE COPY

Applicants: Graham P. Allaway et al.

Serial No.: 09/904,356 Filed: July 12, 2001

Exhibit 14



WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5:		(11) International Publication Number: WO 94/22477					
A61K 39/12, 39/00, C07K 17/00, 3/00, 13/00, 15/00, C07H 15/12, 13/00	A1	(43) International Publication Date: 13 October 1994 (13.10.94)					
(21) International Application Number: PCT/US94/03282 (22) International Filing Date: 25 March 1994 (25.03.94)		(81) Designated States: AU, BR, CA, CN, JP, RU, UA, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).					
(30) Priority Data: 08/037,816 26 March 1993 (26.03.93)	Ũ	Published S With international search report.					
(60) Parent Application or Grant (63) Related by Continuation US 08/037, Filed on 26 March 1993 (
(71) Applicant (for all designated States except US): PRO PHARMACEUTICALS, INC. [US/US]; Old Saw N Road, Tarrytown, NY 10591 (US).	XGENIC VIII Riv	S T					
(72) Inventors; and (75) Inventors/Applicants (for US only): HASEL, I [CH/US]: 1713 Baldwin Road, Yorktown Hei 10598 (US). MADDON, Paul, J. [US/US]: A #25C, 60 Haven Avenue, New York, NY 10032 (ghts, N Apartme	Y					
(74) Agent: WHITE, John, P.; Cooper & Dunham, 30 Re Plaza, New York, NY 10112 (US).	xkefelle						
(54) Title: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF							
(57) Abstract							
The invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein, vaccines comprising the mutant HIV-1 envelope glycoprotein, antibodies and methods of treating individuals.							
		·					

o ilicants: Graham P. Allaway et al. if al No.: 09/904,356 if t: July 12, 2001 is thit 14

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

ΑT	Austria	GB	United Kingdom	MIR	Mauritania
ΑÜ	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	BU	Hungary	NO	Norway
BG	Bulgeria	TE.	Ireland	NZ	New Zealand
BJ	Benin	n	Italy	PL	Poland
BR	Brazil	JP ·	Japan	PT	Portugal
BY	Belane	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic	SD	Sudan
CG	Congo		of Korea	SE	Sweden
CB	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	K2	Kazakhsian	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	.Senegal
CN	China	LK	Sri Lanka	TD	Chad
cs	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobego
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	· US	United States of America
FI	Finland	MIL	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Vict Nam
			•		

HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF

Background of the Invention

10

Throughout this application, various publications are referenced by Arabic numerals. Full citations for these references may be found at the end of the specification immediately preceding the claims. The disclosure of these publications is hereby incorporated by reference into this application to describe more fully the art to which this invention pertains.

The life cycle of animal viruses is characterized by a series of events that are required for the productive infection of the host cell. The initial step in the replicative cycle is the attachment of the virus to the cell surface, which attachment is mediated by the specific interaction of the viral attachment protein (VAP) to 25 receptors on the surface of the target cell. differential pattern of expression of these receptors is largely responsible for the host range and tropic properties In addition, an effective immune response of viruses. against many viruses is mediated through neutralizing 30 antibodies directed against the VAP. The interaction of the VAP with cellular receptors and the immune system therefore plays a critical role in infection and pathogenesis of viral disease.

35 The human immunodeficiency virus type 1 (HIV-1) infects primarily helper T lymphocytes, dendritic cells, and monocytes/macrophages--cells that express surface CD4--leading to a gradual loss of immune function. This loss of function results in the development of the human acquired

2

immunodeficiency syndrome (AIDS) (1). The initial phase of the HIV-1 replicative cycle involves the high-affinity interaction between the HIV-1 exterior envelope glycoprotein gp120 and cell surface CD4 (K_d approximately 4 x 10.9 M) (2). 5 Several lines of evidence demonstrate the requirement of this interaction for viral infectivity. The introduction into CD4 human cells of cDNA encoding CD4 is sufficient to render otherwise resistant cells susceptible to HIV-1 infection (3). In vivo, viral infection appears to be 10 restricted to cells expressing CD4, indicating that the cellular tropism of HIV-1 is largely determined by the pattern of cellular expression of CD4. binding of HIV-1 gp120 to cell surface CD4, viral and target cell membranes fuse by a mechanism that is poorly 15 understood, resulting in the introduction of the viral capsid into the target cell cytoplasm (4).

Mature CD4 has a relative molecular mass (Mr) of 55 kDa and consists of an N-terminal 372-amino acid extracellular 20 domain containing four tandem immunoglobulin-like regions (V1-V4), followed by a 23-amino acid transmembrane domain and a 38-amino acid cytoplasmic segment (5, 6). experiments using truncated sCD4 proteins, it has been shown that the determinants for high-affinity binding to HIV-1 gp120 lie solely within the N-terminal immunoglobulin-like domain (V1) (7-9). Mutational analysis of V1 has defined a discrete binding site (residues 38-52) that comprises a structurally homologous to the second complementarity-determining region (CDR2) of immunoglobulin genes (9). 30

The production of large quantities of sCD4 has permitted a structural analysis of the two N-terminal immunoglobulin-like domains (V1V2). The structure determined at 2.3

3

angstrom resolution reveals that the molecule has two tightly-associated domains, each of which contains the immunoglobulin-fold connected by a continuous beta strand. The putative binding sites for monoclonal antibodies, class II major histocompatibility complex (MHC) molecules, and HIV-1 gp120, as determined by mutational analyses, map on the molecular surface (10, 11).

The HIV-1 envelope gene env encodes an envelope glycoprotein precursor, gp160, which is cleaved by cellular proteases before transport to the plasma membrane to yield gp120 and gp41. The membrane-spanning glycoprotein, gp41, is non-covalently associated with gp120, a purely extracellular glycoprotein. The mature gp120 molecule is heavily glycosylated (approximately 24 N-linked oligosaccharides), contains approximately 480 amino acid residues with 9 intrachain disulfide bonds (12), and projects from the viral membrane as a dimeric or multimeric molecule (13).

Mutational studies of HIV-1 gp120 have delineated important functional regions of the molecule. The regions of gp120 that interact with gp41 map primarily to the N- and C-termini (14). The predominant strain-specific neutralizing epitope on gp120 is located in the 32-34 amino acid residue third variable loop, herein referred to as the V3 loop, which resides near the center of the gp120 sequence (15). The CD4 binding site maps to discontinuous regions of gp120 that include highly conserved or invariant amino acid residues in the second, third, and fourth conserved domains (the C2, C3, and C4 domains) of gp120 (16). It has been postulated that a small pocket formed by these conserved residues within gp120 could accommodate the CDR2 loop of CD4, a region defined by mutational analyses as important in interacting with gp120 (17).

4

HIV-1 gp120 not only mediates viral attachment to surface CD4 molecules, but also serves as the major target of antibodies which neutralize non-cell-associated virus and inhibit cell to cell viral transmission.

5

There are two major classifications of HIV-1-neutralizing antibodies: type-specific and group-common (15). specific neutralizing antibodies primarily recognize linear determinants in the highly variable V3 loop of gp120. These 10 antibodies act by inhibiting fusion between HIV-1 and the target cell membrane, and generally neutralize only a particular isolate of, or closely related strains of, HIV-1. Sequence variation within the V3 loop, as well as outside of this region, permits viruses to escape neutralization by 15 anti-V3 loop antibodies. In contrast, group-common neutralizing antibodies primarily recognize discontinuous or conformational epitopes in gpl20, and possess the ability to neutralize a diverse range of HIV-1 isolates. These broadly neutralizing antibodies often recognize a site on gp120 20 which overlaps the highly conserved CD4-binding site, and thus inhibits gp120-CD4 binding.

A structural relationship has been demonstrated between the V3 loop and the C4 region of gp120 which region constitutes both part of the CD4 binding site and part of the conserved neutralization epitopes. It was observed that deleting the V3 loop resulted in significantly increased binding of a panel of broadly neutralizing hMoAbs (neutralizing human monoclonal antibodies) to the CD4 binding site (18).

30

A major goal in AIDS vaccine development is to develop a vaccine able to protect a subject against the numerous genetic variants of HIV-1 that infect humans. Although cell-mediated immune responses might serve to control infection in HIV-1-infected individuals, several lines of

5

evidence demonstrate that protection against infection is mainly mediated by neutralizing antibodies directed against Early experiments showed that immunization of chimpanzees with recombinant gp120 induced a protective 5 immune response against challenge with the homologous HIV-1 strain (17). This protection correlated with the presence of high-titer neutralizing antibodies against the V3 loop of In addition, passive immunization of chimpanzees with a V3-loop neutralizing monoclonal antibody resulted in 10 protection against challenge with the homologous HIV-1 strain (19). Although protection against challenge was demonstrated in these two experiments, recent studies have questioned the clinical relevance of these findings. example, these neutralizing antibodies recognize the V3 loop 15 determinants of a single strain, and not conserved or discontinuous epitopes. Thus, these antibodies lack the ability to neutralize the broad spectrum of HIV-1 strains present in an HIV-1 population. Furthermore, the challenge virus was the homologous HIV-1 laboratory adapted LAI (HTLV-20 IIIB) strain and not one of the primary isolates that contain considerable gp120 sequence heterogeneity. these experiments showed that gp120 subunit vaccination induces an immune response effective against only the homogeneous HIV-1 strain used as an antigen, it is unlikely 25 that the vaccination regimens used in these studies would be useful in humans.

Individuals infected by HIV-1 typically develop antibodies that neutralize the virus in vitro, and neutralization titers decrease with disease progression (19). Analysis of sera from HIV-1-infected humans indicates that type-specific neutralizing antibodies appear early in infection. Later in the course of infection, a more broadly neutralizing antibody response develops. However this antibody response is of significantly lower titer and/or affinity.

Fractionation studies of HIV-1 antibody-positive human sera reveal that the type-specific neutralizing activity is primarily directed against linear determinants in the V3 loop of gp120 (20). There was no correlation found among antibodies between the ability to neutralize divergent HIV-1 isolates and reactivity to the V3 loop of these isolates. In contrast, the broadly neutralizing antibodies present in HIV-1 antibody-positive human sera primarly recognize discontinuous epitopes in gp120 which overlap the CD4-binding site and block gp120-CD4 binding. In other words, the broadly neutralizing activity of neutralizing antibodies is not merely the result of additive anti-V3 loop reactivities against diverse HIV-1 isolates which appear during infection.

15

Recently, several groups have generated human monoclonal antibodies (hMoAbs) derived from HTV-1 infected individuals which possess type-specific or group-common neutralizing activities (17). The type-specific neutralizing hMoAbs were found to recognize linear determinants in the V3 loop of gp120. In contrast, the group-common neutralizing hMoAbs generally recognize discontinuous epitopes which overlap the CD4-binding site and block gp120-CD4 binding.

The V3 loop is a highly immunodominant region of gp120 which partially interacts with the CD4-binding region. The presence of the V3 loop region on gp120 may skew the humoral immune response away from producing antibodies which specifically bind to the CD4-binding domain of gp120.

Furthermore, the advantages of removing the V3 loop to expose the CD4-binding domain of gp120 to the immune system would be countered by the fact that the exposed CD4-binding site would still have a high affinity for cell surface CD4. In other words, a mutant gp120 protein missing only the V3 loop would quickly bind to CD4+ cells and would thus be

7

hampered in generating an immune response against the exposed CD4-binding site.

The subject invention provides a mutant HIV-1 gp120 envelope glycoprotein which overcomes both the problems of V3 loop immunodominance and of the high affinity to CD4. The subject invention further provides vaccines comprising the mutant HIV-1 gp120 envelope glycoprotein, antibodies which specifically bind to the CD4-binding site of HIV-1 gp120 envelope glycoprotein, pharmaceutical compositions comprising these antibodies, and methods of using these vaccines and compositions to treat or prevent HIV-1 infection.

Summary of the Invention

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W...) point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- 10 In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- 15 In one embodiment, the C4 domain is an HIV-1_{LAI} gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1_{LAI} gp120 envelope glycoprotein.
- In another embodiment, the C4 domain is an HIV-1 $_{\rm IR-FL}$ gp120 envelope glycoprotein C4 domain. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1 $_{\rm IR-FL}$ gp120 envelope glycoprotein.
- 25 The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.
- The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.

The subject invention further provides a method of treating

9

an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.

5 The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.

The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

In one embodiment, the subject is a medical practitioner. In another embodiment, the subject is a newborn infant.

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the subject is a medical practitioner.

12

Brief Description of the Figures

Figure 1

gp120 structure. Shown is a box diagram of HIV-1 gp120 depicting the boundaries of the five constant domains (C1-C5) and the five variable domains (V1-V5). The amino acid residue numbering above the box begins at the initiator methionine found at the beginning of the signal sequence (S) and is approximated based on a consensus of all known HIV-1 gp120 amino acid sequences. Also shown are the C4 domain amino acid sequences of HIV-1 strains LAI and JR-FL. Above the C4 domain sequences are indicated two mutations that reduce gp120 binding to cell surface CD4; tryptophan to valine and aspartate to alanine.

15

Figure 2

PPI4-tPA-gp120LAI. Expression vector with the HIV-1LAI gp120
gene fused to the CMV MIE promoter, and the tPA signal
sequence replacing the HIV-1 gp120 signal sequence.
20 Abbreviations: CMV MIE = cytomegalovirus major immediate
early, E = enhancer, P = promoter, EXA = Exon A, INA =
Intron A, EXB = Exon B, tPA ss = human tissue plasminogen
activator signal sequence, gp120 = glycoprotein 120, BGH =
bovine growth hormone, AMP = ampicillin resistance gene, and
DHFR = dihydrofolate reductase gene.

Figure 3

CMV MIE promoter fused to tPA-gp120_{LAI}. The nucleotide sequence of the CMV MIE promoter/enhancer region is shown fused to the HIV-1_{LAI} gp120 gene that contains the tPA signal sequence. The numbering of nucleotide sequence begins with the HincII site and the numbering of the amino acid sequence begins with the first methionine found in the tPA signal sequence. The tPA signal sequence is fused in-frame to Thr₃₁

13

of gp120, the first amino acid found in mature gp120. The signal sequence is shown in bold as are various landmark restriction sites used for cloning as discussed in the text. The locations of Exon A, Intron A, Exon B and the transcription start site and the signal cleavage site are indicated.

Figure 4

Transient expression of gp120. Autoradiograph of ³⁵S-labeled supernatants from COS cell transfectants, immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. The plasmids used for transfection were: Lane 1: Mock transfected cells; lane 2: a vector encoding a CD4-immunoglobulin chimera as a positive transfection control; lane 3: PPI4-tPA-gp120_{IAI}; and lane 4: PPI4-tPA-gp120_{IR-FL}. Positions of molecular weight markers are indicated.

Figure 5

Determination of gp120 concentration by ELISA. Panel A:
Concentrations of gp120 in media of CHO cell lines, stably
transfected with PPI4-tPA-gp120_{LAI}, determined by ELISA.
Panel B: A standard curve was established using known
amounts of gp120.

25

Figure 6

Expression of gp120 in stably transfected CHO cells.

Autoradiograph of ³⁵S-labeled supernatants from stable CHO cell lines, immunoprecipitated with a CD4-immunoglobulin
Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Lane 1: clone 9; lane 2: clone 13; lane 3: clone 6; lane 4: Clone 5. Positions of molecular weight markers are indicated.

14

Figure 7

tPA-gp120_{JR-FL}. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120 is shown. The NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

Figure 8

tPA-gp120_{LAI}-V3⁽⁾. The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{LAI} gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Thr₃₆ is indicated.

Figure 9

tPA-gp120_{JR-FL}-V3^(*). The nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120 with the V3 loop deleted and replaced with the pentapeptide TGAGH is shown. The V3 loop replacement and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

25

Figure 10

<u>tPA-gp120_{LAI}-V3⁽⁾-CD4⁽⁾</u>. Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV- 1_{LAI} gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp₄₀₈ mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Thr₃₆ is indicated.

15

Figure 11

tPA-gp120_{JR-FL}-V3^(*)-CD4^(*). Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{JR-FL} gp120, with the V3 loop deleted and replaced with the pentapeptide TGAGH, and Trp₃₉₆ mutated to Val. The mutations and the NarI and NotI restriction endonuclease sites used for cloning are shown in bold. The predicted site of cleavage by signal peptidase between Arg₃₅ and Val₃₆ is indicated.

10

Figure 12

tPA-qp120_{LAI}-CD4^(*). Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{LAI} gp120. The Trp₄₃₇ to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning, and the predicted site of cleavage by signal peptidase between Arg₃₅ and Thr₃₆ are shown in bold.

Figure 13

tPA-gp120_{R-FL}-CD4^(*). Shown is the nucleotide and deduced amino acid sequence of the tPA signal sequence fused to HIV-1_{IR-FL} gp120. The Trp₄₂₄ to Val CD4 binding mutation, the NarI and NotI restriction endonuclease sites used for cloning and the predicted cleavage by signal peptidase between Arg₃₅ and Val₃₆ are shown in bold.

Figure 14

Expression of gp120 in stably transfected CHO cells.

Autoradiograph of super ³⁵S-labeled supernatants from stable CHO cell lines, immunoprecipitated with MoAb F105-Protein A-Sepharose complex, and run on a reducing 10% SDS-PAGE gel. Panel A: Lane 1: tPA-gp120_{LAI} CHO cells; lane 2: tPA-gp120_{LAI}-V3⁽⁴⁾ CHO cells; lane 3: tPA-gp120_{LAI}-V3⁽⁴⁾-CD4⁽⁴⁾ CHO cells. Panel B: Lane 1: tPA-gp120_{IRFL} CHO cells; lane 2: tPA-gp120_{IRFL}-V3⁽⁴⁾

CHO cells; lane 3: $tPA-gp120_{JR-FL}-V3^{(\cdot)}-CD4^{(\cdot)}$ CHO cells. Positions of molecular weight markers are indicated.

Figure 15

5 Purified ap120 proteins.

Silver stained 10% SDS-PAGE gel with a sample of purified gp120 proteins. Panel A: Lane 1: tPA-gp120_{LAI} CHO cells; lane 2: tPA-gp120_{LAI}-V3^(·) CHO cells; lane 3: tPA-gp120_{LAI}-V3^(·)-CD4^(·) CHO cells. Panel B: Lane 1: tPA-gp120_{IR-FL} CHO cells; lane 2: tPA-gp120_{IR-FL}-V3^(·) CHO cells; lane 3: tPA-gp120_{IR-FL}-V3^(·)-CD4^(·) CHO cells. Positions of molecular weight markers are indicated.

Figure 16

Analysis of binding of recombinant mutant gp120 to cell 15 surface human CD4 by FACS.

Plate 1. DG44 cells, a subclone of CHO cells which lack expression of the human CD4 protein, were used as control. Increasing concentrations of HIV-1 gp120LAI did not show an specific fluoresence when compared to increase in Plate 2. DG44 #3 cells are a CHO cell line 20 background. transfected with the cDNA clone encoding the human CD4 protein. Increasing concentrations of HIV-1 gp120LAI show a dramatic increase (or shift) in fluoresence. Similar to Plate 2 but the HIV-1 gp120LAI-V3(-) protein was 25 added. Again a large shift indicating binding to the DG44 #3 cells was seen. Plate 4. DG44 #3 cells were incubated with either HIV-1 gp120_{LAI}-V3^(·)-CD4^(·) protein or MoAb OKT4A an antibody with high affinity for human CD4. Only OKT4A bound to the cells.

17

Detailed Description of the Invention

The plasmids designated PPI4-tPA-gp120_{LAI} and PPI4-tPA-gp120_{JR}.

FL were deposited pursuant to, and in satisfaction of, the

requirements of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, Maryland 20852 under ATCC Accession Nos. 75431 and 75432, respectively. The plasmids PPI4-tPA-gp120_{LAI} and PPI4-tPA-gp120_{IR-FL} were deposited with the ATCC on March 12, 1993.

The subject invention provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation, wherein X is an amino acid residue other than tryptophan. In the preferred embodiment, X is a valine residue.

- In one embodiment, the nucleic acid molecule is a DNA molecule. The DNA molecule may be a plasmid. In one embodiment, the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
- The V3 loop of HIV-1 gp120 envelope glycoprotein is shown in Figure 1. The V3 loop is demarcated by cysteine residues at both its N- and C-termini. As used herein, a V3 loop deletion means a deletion of one or more amino acid residues between the terminal cysteine residues, with the proviso that there must be three or more amino acid residues situated between the two terminal cysteine residues in a V3 loop deletion. These three or more amino acid residues may either be residues originally present in the V3 loop, or exogenous residues. For example, as shown in the

WO 94/22477

Experimental Details section <u>infra</u>, the pentapeptide TGAGH is situated between the two terminal cysteine residues. Variations in the size of the V3 loop deletion illustrated herein are tolerable without affecting the overall structure of the mutant HIV-1 gp120 envelope glycoprotein, as is well known to those skilled in the art.

As used herein, "C4 domain" means the HIV-1 gp120 envelope glycoprotein C4 domain having the following consensus 10 sequence:

 $\begin{array}{l} X_{1} X_{2} X_{3} C X_{4} I X_{5} X_{6} X_{7} X_{8} X_{9} X_{10} W X_{11} X_{12} X_{13} X_{14} X_{15} A X_{16} Y X_{17} X_{18} - \\ P X_{19} X_{20} X_{21} X_{22} X_{23} X_{24} X_{25} X_{26} S X_{27} X_{28} TG X_{29} X_{30} X_{31} X_{32} R X_{33} G X_{34}, \end{array}$

15 wherein X₁ = T, I, V, K or R; X₂ = L, I or H; X₃ = P, Q, L or
T; X₄ = R, K or G; X₅ = K or E; X₆ = Q or E; X₇ = F, I or V;
X₈ = I, V or M; X₉ = N, R or K; X₁₀ = M, R, L or T; X₁₁ = Q, R
or V; X₁₂ = E, K, G, R, V or A; X₁₃ = V, T, A or G; X₁₄ = G or
E; X₁₅ = K, R, E, or Q; X₁₆ = M, V, I or L; X₁₇ = A, T or D; X₁₈
20 = P or L; X₁₉ = I or F; X₂₀ = S, R, G, K, N, A, E or Q; X₂₁ =
G or R; X₂₂ = Q, L, P, N, K, V, T, E or I; X₂₃ = I, V or L; X₂₄
= R, K, S, N, G, I, T, E or I; X₂₅ = C or R; X₂₆ = S, L, I, T,
P, E, V, K, D or N; X₂₇ = N, K or L; X₂₈ = I or V; X₂₉ = L, P
or I; X₃₀ = L or I; X₃₁ = L or I; X₃₂ = T, A, I, V or E; X₃₃ =
25 D or E; X₃₄ = G or V.

The C4 domain consensus sequence is based on existing C4 domain sequence information from various HIV-1 strains, and thus is not necessarily an exhaustive consensus sequence. The conserved tryptophan residue shown in bold after residue X_{10} is the only conserved tryptophan residue in the C4 domain. As used herein, a C4 domain $_{(W->X)}$ point mutation is a mutation of the above-identified conserved C4 domain tryptophan residue to an amino acid residue other than

19

tryptophan. For example, a C4 domain $_{(W->V)}$ point mutation is a mutation of the conserved C4 domain tryptophan residue to a valine residue.

In one embodiment, the C4 domain is an HIV-1_{LAI} gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1 LAI 9P120 C4 domain is: TLPCRIKQFINMWQEVGKAMYAPPISGOIRCS-SNITGLLLTRDGG. The mutant HIV-1 gp120 envelope glycoprotein may be a mutant HIV-1LAI gp120 envelope glycoprotein.

10

In another embodiment, the C4 domain is an $HIV-1_{JR-FL}$ gp120 envelope glycoprotein C4 domain. The sequence of the HIV-1, p. FL gp120 C4 domain is: TLPCRIKQIINMWQEVGKAMYAPPIRGQIRCS-SNITGLLLTRDGG. The mutant HTV-1 gp120 envelope glycoprotein

may be a mutant $HIV-1_{JR-FL}$ gp120 envelope glycoprotein.

HIV-1LAI is a laboratory-adapted strain that is tropic for phytohemagglutinin (PHA) -stimulated peripheral lymphocytes (PBLs) and immortalized human T-cell lines. In 20 contrast, HIV-1_{R-FL} was isolated from brain tissue taken at autopsy that was co-cultured with lectin-activated normal human PBLs. HIV-1_{IR-FL} is tropic for PHA-stimulated PBLs and blood-derived macrophages but will not replicate in transformed T-cell lines. Mutant HIV-1 gp120 envelope 25 glycoproteins derived from a clinical isolate of HIV-1 such as JR-FL may possess new or different epitopes compared to the laboratory-adapted HIV-1 strains that are beneficial for successful vaccination. Although only the HIV-1141 and HIV- $\mathbf{1}_{\text{IR-FL}}$ strains are used herein to generate the mutant HIV-1 30 gp120 envelope glycoproteins of the subject invention, other HIV-1 strain could be substituted in their place as is well known to those skilled in the art.

The V1 and V2 variable regions of gp120 are unnecessary for

CD4 binding (21). Therefore the mutant HIV-1 gp120 envelope glycoprotein of this invention can either include or exclude the V1 and V2 variable regions.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gpl20 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(Asp->X) point mutation, wherein the aspartate residue is between amino acid residues X₁₅ and X₁₆ in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain $_{(Gh->X)}$ point mutation, wherein the glutamate residue is between amino acid residues X_{15} and X_{16} in the C4 consensus sequence, and X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is an alanine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{LAI} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_(asp378->X) point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{R-FL} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_(asp369->X) point mutation, wherein X is an amino acid residue other than aspartate or glutamate. In the preferred

21

embodiment, X is a lysine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{LAI} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_{gh380->X)} point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment, X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{JR-FL} gp120 envelope glycoprotein comprising a V3 loop deletion and a C3 domain_(ghi371->X) point mutation, wherein X is an amino acid residue other than glutamate. In the preferred embodiment,

15 X is a glutamine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{LAI} gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain_(th/267->X) point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1_{IR-FL} gp120 envelope glycoprotein comprising a V3 loop deletion and a C2 domain_(th/260->X) point mutation, wherein X is an amino acid residue other than threonine. In the preferred embodiment, X is an arginine residue.

30

The subject invention additionally provides a recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising (a) a V3 loop deletion, or (b) a one of the C2, C3 or C4 domain point mutations

discussed supra.

The point mutations in the recombinant nucleic acid molecules described <u>supra</u> are selected based on their ability to reduce the affinity of the mutant gp120 glycoprotein encoded thereby for CD4. As used herein, the term "reduce the affinity" means to reduce the affinity by at least two-fold.

One skilled in the art would know how to make recombinant nucleic acid molecules which encode mutant HIV-1 gp120 envelope glycoproteins comprising a V3 loop deletion and the specific C2, C3 or C4 domain point mutations corresponding to those mutations exemplified in the HIV-1_{R-FL} and HIV-1_{LA1} strains, supra. Furthermore, one skilled in the art would know how to use these recombinant nucleic acid molecules to obtain the proteins encoded thereby, and practice the therapeutic and prophylactic methods of using same, as described herein for the recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a V3 loop deletion and a C4 domain_(W->X) point mutation.

The subject invention also provides the mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of the subject invention.

In accordance with the invention, numerous vector systems for expression of the mutant HIV-1 gp120 envelope glycoprotein may be employed. For example, one class of vectors utilizes DNA elements which are derived from animal viruses such as bovine papilloma virus, polyoma virus, adenovirus, vaccinia virus, baculovirus, retroviruses (RSV, MMTV or MoMLV), Semliki Forest virus or SV40 virus.

Additionally, cells which have stably integrated the DNA into their chromosomes may be selected by introducing one or more markers which allow for the selection of transfected host cells. The marker may provide, for example, prototropy auxotrophic host, biocide resistance, antibiotics) or resistance to heavy metals such as copper or the like. The selectable marker gene can be either directly linked to the DNA sequences to be expressed, or introduced into the same cell by cotransformation. Additional elements 10 may also be needed for optimal synthesis of mRNA. elements may include splice signals, enhancers, and termination transcriptional promoters, The cDNA expression vectors incorporating such elements include those described by Okayama (22).

15

The vectors used in the subject invention are designed to express high levels of mutant HIV-1 gp120 envelope glycoproteins in cultured eukaryotic cells as well as efficiently secrete these proteins into the culture medium.

The targeting of the mutant HIV-1 gp120 envelope glycoproteins into the culture medium is accomplished by fusing in-frame to the mature N-terminus of the mutant HIV-1 gp120 envelope glycoprotein the tissue plasminogen activator (tPA) prepro-signal sequence.

25

The mutant HIV-1 gp120 envelope glycoprotein may be produced by a) transfecting a mammalian cell with an expression vector for producing mutant HIV-1 gp120 envelope glycoprotein; b) culturing the resulting transfected mammalian cell under conditions such that mutant HIV-1 gp120 envelope glycoprotein is produced; and c) recovering the mutant HIV-1 gp120 envelope glycoprotein so produced.

Once the expression vector or DNA sequence containing the constructs has been prepared for expression, the expression

vectors may be transfected or introduced into an appropriate mammalian cell host. Various techniques may be employed to achieve this, such as, for example, protoplast fusion, calcium phosphate precipitation, electroporation or other conventional techniques. In the case of protoplast fusion, the cells are grown in media and screened for the appropriate activity. Expression of the gene encoding a mutant HIV-1 gp120 envelope glycoprotein results in production of the mutant glycoprotein.

24

10

Methods and conditions for culturing the resulting transfected cells and for recovering the mutant HIV-1 gp120 envelope glycoprotein so produced are well known to those skilled in the art, and may be varied or optimized depending upon the specific expression vector and mammalian host cell employed.

In accordance with the claimed invention, the preferred host cells for expressing the mutant HIV-1 gp120 envelope glycoprotein of this invention are mammalian cell lines. Mammalian cell lines include, for example, monkey kidney CV1 line transformed by SV40 (COS-7); human embryonic kidney line 293; baby hamster kidney cells (BHK); Chinese hamster ovary-cells-DHFR (CHO); Chinese hamster ovary-cells DHFR (DXB11); monkey kidney cells (CV1); African green monkey kidney cells (VERO-76); human cervical carcinoma cells (HELA); canine kidney cells (MDCK); human lung cells (W138); human liver cells (Hep G2); mouse mammary tumor (MMT 060562); mouse cell line (C127); and myeloma cell lines.

30

Other eukaryotic expression systems utilizing non-mammalian vector/cell line combinations can be used to produce the mutant HIV-1 gp120 envelope glycoproteins. These include, but are not limited to, baculovirus vector/insect cell

expression systems and yeast shuttle vector/yeast cell expression systems.

Methods and conditions for purifying mutant HIV-1 gp120 envelope glycoproteins from the culture media are provided in the invention, but it should be recognized that these procedures can be varied or optimized as is well known to those skilled in the art.

- 10 The subject invention further provides a vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 15 A therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- As used herein, adjuvants include, but are not limited to, alum, Freund's incomplete adjuvant (FIA), Saponin, Quil A, Monophosphoryl lipid A (MPL), and nonionic block copolymers (SAF) such as L-121 (Pluronic; Syntex SAF). In the preferred embodiment, the adjuvant is alum, especially in the form of a thixotropic, viscous, and homogeneous aluminum hydroxide gel. The vaccine of the subject invention may be administered as an oil in water emulsion. Methods of combining adjuvants with antigens are well known to those skilled in the art.
- 30 The subject invention further provides a method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of the subject invention, thereby treating the HIV-1-infected subject.
- 35 As used herein, treating an HIV-1-infected subject with the

vaccine of the subject invention means reducing in the subject either the population of HIV-1 or HIV-1-infected cells, or ameliorating the progression of an HIV-1-related disorder in the subject.

5

As used herein, an "HIV-infected subject" means an individual having at least one of his own cells invaded by HIV-1.

- As used herein, "immunizing" means administering a primary dose of the vaccine to a subject, followed after a suitable period of time by one or more subsequent administrations of the vaccine, so as to generate in the subject an immune response against the CD4-binding region of the mutant HIV-1 gp120 envelope glycoprotein in the vaccine. A suitable period of time between administrations of the vaccine may readily be determined by one skilled in the art, and is usually in the order of several weeks to months.
- 20 In the preferred embodiment, the dose of vaccine administered is an amount sufficient to deliver to the subject between 10ug and 1mg of the mutant HIV-1 gp120 envelope glycoprotein.
- The subject invention further provides a vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of the subject invention, and an adjuvant.
- 30 A prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein may be determined according to methods well known to those skilled in the art.
- The subject invention further provides a method of reducing the likelihood of an HIV-1-exposed subject's becoming

27

infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

5

As used herein, the subject's becoming infected with HIV-1 means the invasion of the subject's own cells by HIV-1.

As used herein, reducing the likelihood of a subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least two-fold. For example, if a subject has a 1% chance of becoming infected with HIV-1, a two-fold reduction in the likelihood of the subject's becoming infected with HIV-1 would result in the subject's having a 0.5% chance of becoming infected with HIV-1. In the preferred embodiment of this invention, reducing the likelihood of the subject's becoming infected with HIV-1 means reducing the likelihood of the subject's becoming infected with HIV-1 by at least ten-fold.

As used herein, an HIV-1-exposed subject is a subject who has HIV-1 present in his body, but has not yet become HIV-1-infected.

25

The subject invention further provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of the subject invention, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

As used herein, a non-HIV-1-exposed subject is a subject who does not have HIV-1 present in his body.

The subject invention further provides a method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of the subject invention, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein. In the preferred embodiment, the subject is a human.

As used herein, partially purified antibodies means a composition which comprises antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein, and consists of fewer protein impurities than does the serum from which the anti-CD4-binding domain antibodies are derived. A protein impurity means a protein other than the anti-CD4-binding domain antibodies. For example, the partially purified antibodies might be an IgG preparation.

Methods of recovering serum from a subject are well known to those skilled in the art. Methods of partially purifying antibodies are also well known to those skilled in the art, and include, by way of example, filtration, ion exchange chromatography, and precipitation.

In one embodiment, the partially purified antibodies comprise an immune globulin (IG) preparation. IG can be purified from serum by a two-step process. Initially, serum is fractionated by the cold ethanol method of Cohn, et al. (29). Cohn Fraction II has as its main protein component IgG immunoglobulin present as monomers, dimers and aggregates. Fraction II is then purified to produce IVIG

29

(immune globulin intravenous) using a variety of purification methods which include, for example, ion exchange, DEAE chromatography, acid pH 4.25 diafiltration, PEG precipitation or Pepsin treatment. The final product is stabilized (e.g., glucose + NaCl) and the final IgG concentration is fixed at between about 3% and about 6%.

The subject invention further provides the partially purified antibodies produced by the method of the subject 10 invention.

The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A therapeutically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

20

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to. 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% Additionally, such pharmaceutically acceptable saline. 25 carriers may be aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous are propylene glycol, polyethylene vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include alcoholic/aqueous solutions, 30 water. emulsions suspensions, including saline and buffered Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid 35 and nutrient replenishers, electrolyte replenishers such as

29

(immune globulin intravenous) using a variety of purification methods which include, for example, ion exchange, DEAE chromatography, acid pH 4.25 diafiltration, PEG precipitation or Pepsin treatment. The final product is stabilized (e.g., glucose + NaCl) and the final IgG concentration is fixed at between about 3% and about 6%.

The subject invention further provides the partially purified antibodies produced by the method of the subject invention.

The subject invention further provides a pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A therapeutically effective amount of the partially purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

20

Pharmaceutically acceptable carriers are well known to those skilled in the art and include, but are not limited to, 0.01-0.1M and preferably 0.05M phosphate buffer or 0.8% saline. Additionally, such pharmaceutically acceptable carriers may agueous be or non-aqueous suspensions, and emulsions. Examples of non-aqueous propylene glycol, polyethylene are vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include 30 water, alcoholic/aqueous solutions, emulsions suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution. Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's or fixed oils. Intravenous vehicles include fluid 35 and nutrient replenishers, electrolyte replenishers such as

those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

30

5

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

As used herein, administering may be effected or performed using any of the various methods known to those skilled in the art. The administering may comprise administering intravenously. The administering may also comprise administering intramuscularly. The administering may further comprise administering subcutaneously.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

30

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-

30

those based on Ringer's dextrose, and the like. Preservatives and other additives may also be present, such as, for example, antimicrobials, antioxidants, chelating agents, inert gases and the like.

5

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.

As used herein, administering may be effected or performed using any of the various methods known to those skilled in the art. The administering may comprise administering intravenously. The administering may also comprise administering intramuscularly. The administering may further comprise administering subcutaneously.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

30

The subject invention further provides a method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1-in the HIV-1-infected subject, thereby treating the HIV-1-

PCT/US94/03282

infected subject.

The dose of the pharmaceutical composition of the subject invention effective to reduce the population of HIV-1 in the HIV-1-infected subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10 mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The subject invention further provides a composition which comprises a prophylactically effective amount of the partially purified antibodies of the subject invention, and a pharmaceutically acceptable carrier.

A prophylactically effective amount of the partially 20 purified antibodies of the subject invention may be determined according to methods well known to those skilled in the art.

The subject invention further provides a method of reducing
the likelihood of an HIV-1-exposed subject's becoming
infected with HIV-1, which comprises administering to the
HIV-1-exposed subject a dose of the composition of the
subject invention effective to reduce the population of HIV1 in the HIV-1-exposed subject, thereby reducing the
likelihood of the subject's becoming infected with HIV-1.

In one embodiment, the subject is a medical practitioner.

The medical practitioner may be a medical practitioner exposed to an HIV-1-containing bodily fluid. As used herein,

the term "medical practitioner" includes, but is in no way

32

limited to, doctors, dentists, surgeons, nurses, medical laboratory assistants, and students in health care programs.

In another embodiment, the subject is a newborn infant. The newborn infant may be a newborn infant born to an HIV-1-infected mother.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 in the HIV-110 exposed subject may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100 mg/kg and 2g/kg of protein if administered intravenously.

The vaccines and pharmaceutical compositions of the subject invention may also ameliorate the progression of an HIV-1-related disorder in a subject to whom the vaccines or pharmaceutical compositions were administered while the subject was either non-HIV-1-exposed or HIV-1-exposed, but not yet HIV-1-infected.

25

Finally, the subject invention provides a method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1 as a result of exposure thereto during an incident wherein there is an increased risk of exposure to HIV-1, which comprises administering to the subject immediately prior to the incident a dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident, thereby reducing the likelihood of the subject's becoming infected with HIV-1. In one embodiment, the

subject is a medical practitioner.

An incident wherein there is an increased risk of exposure to HIV-1 includes, for example, receiving a blood transfusion, sexual contact with an HIV-1-infected individual, and performing a HIV-1-containing bodily fluid-exposing medical procedure.

As used herein, "immediately prior to the incident" means 10 within one month of the incident. In the preferred embodiment, "immediately prior to the incident" means within one day of the incident.

The dose of the composition of the subject invention effective to reduce the population of HIV-1 to which the subject is exposed during the incident may be readily determined using methods well known to those skilled in the art. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 10mg/kg and 150mg/kg of protein if administered intramuscularly. In the preferred embodiment, the dose is sufficient to deliver to the subject between about 100mg/kg and 2g/kg of protein if administered intravenously.

this invention is a method embodiment of 25 substantially reducing the likelihood of a non-infected medical practitioner's becoming infected with HTV-1 during a bodily fluid-exposing medical procedure involving a patient, which comprises administering to the patient during a suitable time period an amount of the composition of the 30 subject invention effective to substantially reduce the likelihood of the non-infected medical practitioner's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid during the medical procedure.

As used herein, a bodily fluid is any fluid which is present in the human body and is capable of containing infectious HIV-1 in an HIV-1-infected patient. Bodily fluids include, but are not limited to, saliva, cerebrospinal fluid, tears, vaginal secretions, urine, alveolar fluid, synovial fluid and pleural fluid.

Another embodiment of this invention is a method of substantially reducing the likelihood of a non-HIV-110 infected newborn infant's becoming infected with HIV-1 prior to or during birth from an HIV-1-infected mother, which comprises administering to the mother prior to birth an amount of the composition of the subject invention effective to substantially reduce the likelihood of the non-HIV-115 infected newborn infant's becoming infected with HIV-1 by virtue of contact with the patient's bodily fluid.

In order to facilitate an understanding of the Experimental Details section which follows, certain frequently occurring methods and/or terms are best described in Maniatis et al. (23).

This invention will be better understood by reference to the Experimental Details which follow, but those skilled in the art will readily appreciate that the specific experiments detailed are only illustrative of the invention as described more fully in the claims which follow thereafter.

Experimental Details

Nomenclature

As used herein, V3^(·) indicates a V3 loop deletion from HIV-1 gp120 envelope glycoprotein. As used herein, CD4^(·) indicates a point mutation in the C4 domain of HIV-1 gp120 envelope glycoprotein which mutation inhibits CD4 binding to the mutant HIV-1 gp120 envelope glycoprotein. The structure of HIV-1 gp120 envelope glycoprotein is shown in Figure 1.

10 Materials and Methods

Construction of PPI4-tPA-gp120 expression vector. An expression vector was constructed that consisted of the immediate early (CMV cytomegalovirus major 15 promoter/enhancer linked to the HIV-1 gene, which gene had its signal sequence replaced by the tPA signal sequence. The CMV MIE promoter/enhancer sequences were derived from pSVCC1 (24) consisting of 1580 base pairs of contiguous DNA that is immediately 5' to the initiator ATG. In sequential 20 order, the functional domains of the CMV promoter are: the promoter/enhancer region; a transcriptional initiator site; exon A (a non-coding exon); intron A; and 17 nucleotides of exon B (non-coding sequences). The viral promoter sequences were ligated to a gene construct consisting of the 25 nucleotide sequences encoding amino acids -35 to -1 of human tPA (25) fused in-frame to HIV-1 amino acids 31 through 515, ending with a TGA stop codon. The construction was performed in two parts. The majority of the CMV promoter could be isolated as a 1560 bp Hinc II/Pst I fragment which 30 was ligated to a Pst I/Not I 1590 bp DNA fragment that contained the remainder of the CMV promoter, the initiator ATG, the tPA signal sequence and the mature HIV-1141 env protein coding sequence.

The latter fragment was assembled using the polymerase chain reaction as follows. Primer 1 (GATCCTGCAGTCACCGTCCTTGACA-CGATGGATGCAATGAAGAGA) and primer 2 (AAGTCTTCTCCTCGGTCTTGT-CTTTTTAACACCCAG) were used to amplify the nucleic acid 5 sequences encoding the tPA signal sequence amino acids -35 to -1 from plasmid pMAM neo-s (Clonetech), thus producing a 150 bp fragment. A second 1440 bp DNA fragment was amplified (TTCAGAAGAGGAGCCAGAACAGAAAAATTGTGGGTC). 3 primer primer 4 (GGAAAAAGCGGCCGCTCATTTTTCTCTCTCTGCACCACTC), and pENV (26) as a template. The PCR fragments were pooled, 10 desalted, and excess primer removed by ultrafiltration through a centricon-100 unit (Amicon). An aliquot of the pooled material was then subjected to a second round of amplification in the presence of primers 1 and 4 to produce a 1590 bp fragment, which was then digested with Pst I and 15 The CMV promoter fragment and the HIV-1, env fragment were then ligated together, and the entire transcription unit subcloned into PPI4, which eukaryotic shuttle vector that contains an ampicillin 20 resistance gene, an SV40 origin of replication and a DHFR gene whose transcription is driven by the ß-globin promoter. The final construct, PPI4-tPA-gp120_{LAI}, is shown in Figure 2.

The expression vector is then used as the prototype vector for the expression of gp120 proteins that are derived from other HIV-1 strains or mutated as described in the methods section. The vector was constructed so that unique Nar I and Not I sites flank the gp120 sequence, thus facilitating the removal of the gp120 gene cassette and the subsequent insertion of other gene cassettes (Figure 2).

- 2. Expression of HIV-1, app120 in mammalian cells.
- a. Transient expression.

CosM5 cells grown in DMEM containing 10% fetal calf serum

were split to 75% confluence. On the following day, the cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120_{LAI} DNA by the standard CaPO₄ (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with ³⁵S-cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed by SDS-10 PAGE under reducing conditions (Figure 4).

b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) were transfected with 20 micrograms of CsCl-purified DNA. Approximately 3-5 15 days post-transfection, cells were placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones were picked. Media was analyzed for gp120 expression by radiolabelling the cells with 35S-20 cysteine for 12-18 hours, followed by precipitation of media using a CD4-immunoglobulin-Protein A-Sepharose complex, followed in turn by SDS-PAGE under reducing conditions The levels of gp120 in the media of these (Figure 6). clones were also quantitated (Figure 5) by ELISA performed The method involves coating 96-well plates 25 as follows. overnight with sheep polyclonal IgG against the highly conserved C-terminus of gp120 (D7234, Aalto Bioreagents). After washing, dilutions of a standard gp120 preparation in cell growth medium, or supernatant from the 30 transfected cells, were incubated for 1 hour. The plates were washed again, and incubated for one hour with a horseradish peroxidase-conjugated anti-gp120 monoclonal Following a final wash, the antibody (9204, DuPont). peroxidase substrate OPD (DuPont) was added and the amount WO 94/22477 PCT/US94/03282

38

of qp120 determined by comparing absorbance of unknowns with Standards were prepared from purified a standard curve. gp120 made in CHO cells, a small quantity of which was obtained from Celltech Ltd. Clones expressing the highest 5 levels were subjected to successive rounds of amplification of the newly introduced DNA sequences in concentrations of methotrexate. Stable CHO cell lines were thus generated which secrete at least 1 microgram/milliliter of HIV-1_{LAI} gp120.

10

30

3. Construction of PPI4-tPA-qp120_{R-FL}

The HIV-1_{LAI} gp120 env nucleotide sequence in PPI4-tPAgp120LAI was replaced by the nucleotide sequence encoding the mature gp120_{R-FI} protein. Using the polymerase reaction, the JR-FL sequences were amplified from pUC112-1 (27) using primer 5 (GATCGGCGCCAGAGTAGAAAGTTGTGGGTCAC) and The PCR fragment was digested with restriction endonucleases Nar I and Not I, and the fragment subcloned in between the Nar I and Not I sites in PPI4-tPA-20 gp120_{IAI} to generate PPI4-tPA-gp120_{IR-FL} (Figure 7).

Transient expression.

CosM5 cells grown in DMEM containing 10% fetal calf serum were split to 75% confluence. On the following day, the 25 cells were transfected for 16-20 hours with 10 micrograms of CsCl-purified PPI4-tPA-gp120_{JR-FL} DNA by the standard CaPO₄ (5) precipitation technique. After transfection, fresh medium was added to the cells. Analysis of the products synthesized 96-120 hours post-transfection was performed by radiolabelling the transfectants with 35-cysteine for 12-18 hours, followed by precipitation of media using a CD4immunoglobulin-Protein A-Sepharose complex, followed by SDS-PAGE under reducing conditions (Figure 4).

WO 94/22477

4. Construction of PPI4-tPA-gp120_[A]-V3⁽⁴⁾.

The V3 loop in tPA-gp120_{LAI} consists of amino acids Cys₃₀₆ In the V3⁽⁻⁾ mutant, the amino acids in through Cys33. between these cysteines are replaced by the pentapeptide 5 sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120_{LAI} is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) primer 7 (CTCGAGCATGCATTCGAAGCTCGCTGATC) as a selection primer. Primer 7 changes a unique Xba I site in the 10 backbone of the parent PPI4 plasmid into a unique BstB I Briefly, the mutagenesis method requires incubating of the parent plasmid with the mutagenic primer and the selection primer, denaturing at 100°C for 3 minutes and then 15 chilling on ice. In the presence of buffered deoxynucleotide triphosphates and T4 DNA polymerase, the primers are allowed to initiate the polymerization of one strand of T4 DNA ligase is used to seal the newly plasmid DNA. synthesized DNA strand to form a covalently closed circle. 20 Hybrid plasmids are then transformed into a MutS strain of E. coli that is deficient in mismatch repair. allowing for the growth of transformed cells, DNA is purified from the cells and digested with the selection restriction endonuclease, in this case Xba I. 25 plasmids are cleaved by Xba I while the mutant plasmid remains resistant to cleavage by virtue of the Xba I to BstB Digested DNA is then used to transform E. I conversion. coli, and colonies harboring the mutant plasmid are picked. Multiple mutagenic primers can be used in a single round of The amino acid sequence of the modified 30 mutagenesis. protein is shown in Figure 8.

5. Construction of PPI4-tPA-gp120_{RFI}-V3(4).

The V3 loop in tPA-gp120_{IR-FL} consists of amino acids Cys,

through Cys₃₂₇. In the V3⁽⁻⁾ mutant, the amino acids in between these cysteines are replaced by the pentapeptide sequence Thr-Gly-Ala-Gly-His. Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the V3 loop sequence in PPI4-tPA-gp120_{JR-FL} is altered using the mutagenic primer 6 (CTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAACATTAGTAGAGC) and primer 7 as a selection primer. The amino acid sequence of the modified protein is shown in Figure 9.

10 6. Construction of PPI4-tPA-gp120_{LAI}-CD4⁽⁻⁾.

Using the Transformer Site-Directed Mutagenesis Kit (Clonetech), the selection primer 7. and the mutagenic primer 8 (CAATTTATAAACATGGTGCAGGAAGTAGG), Trp₄₃₇ of tpA-gp120_{LAI}, which is in an equivalent position to the tryptophan residue in the HXBc2 strain of HIV-1, is mutated to a Val in the expression vector PPI4-tpA-gp120_{LAI} to generate PPI4-tpA-gp120_{LAI}-CD4⁽⁻⁾. The sequence for gp120_{LAI}-CD4⁽⁻⁾ is shown in Figure 12.

20 7. Construction of PPI4-tPA-qp120_{RFI}-CD4(4).

In a fashion similar to that described above, Trp_{424} of $tpA-gp120_{IR-FL}$ is mutated to a Val in the expression vector PPI4- $tPA-gp120_{IR-FL}$ using the selection primer 7 and the mutagenic primer 9 (CAAATTATAAACATGGTGCAGGAAGTAGG) to generate PPI4- $tPA-gp120_{IR-FL}-CD4^{(\cdot)}$. The sequence for $gp120_{IR-FL}-CD4^{(\cdot)}$ is shown in Figure 13.

8. Construction of PPI4-tPA-gp120_{LAI}-V3⁽⁾-CD4⁽⁾.

The tPA-gp120_{LAI} double mutant, V3⁽⁴⁾-CD4⁽⁴⁾, is constructed by including the mutagenic primers 6 and 8, and the selection primer 7 simultaneously in the reaction tube with PPI4-tPA-gp120_{LAI} as the DNA template. The final construct is named PPI4-tPA-gp120_{LAI}-V3⁽⁴⁾-CD4⁽⁴⁾, and its sequence is shown in figure 10.

9. Construction of PPI4-tPA-gp120_{JR-FL}-V3⁽⁻⁾-CD4⁽⁻⁾.

The tPA-gp120_{JR-FL} double mutant, V3^(·)-CD4^(·), is constructed by including the mutagenic primers 6 and 9, and the selection primer 7 simultaneously in the reaction tube with PPI4-tpA-gp120_{JR-FL} as the DNA template. The final construct is named PPI4-tPA-gp120_{JR-FL}-V3^(·)-CD4^(·), and its sequence is shown in figure 11.

10. Expression of mutant HIV-1 gp120 in mammalian cells.

10 a. <u>Transient expression</u>.

CosM5 cells grown in DMEM containing 10% fetal calf serum are split to 75% confluence. On the next day, the cells are transfected for 16-20 hours with 10 micrograms of CsCl-purified mutant HIV-1 DNA by the standard CaPO₄ (5) precipitation technique. After transfection, fresh medium is added to the cells. Analysis of the products synthesized 96-120 hours post-transfection is performed by radiolabelling the transfectants with 35S-cysteine for 12-18 hours, followed by precipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120.

b. Stable expression.

Dhfr Chinese hamster ovary cells (CHO) are transfected with 20 micrograms of CsCl-purified DNA encoding the native or mutant HIV-1 gp120 glycoproteins. Approximately 3-5 days post-transfection, cells are placed in selective medium (nucleoside-free alpha MEM containing 10% dialyzed fetal calf serum). Approximately 10-15 days post-selection, individual cell clones are picked. Media is analyzed for gp120 expression by radiolabelling the cells with 35s-cysteine for 12-18 hours, followed by quantitative immunoprecipitation of media using a sheep polyclonal IgG against the highly conserved C-terminus of gp120, followed

SDS - PAGE turn by under reducing conditions. Alternatively, one can quantitate the level of gp120 by ELISA performed as follows. The method involves coating 96well plates overnight with sheep polyclonal IgG against the 5 highly conserved C-terminus of gp120 (D7234, Bioreagents). After washing, dilutions of a standard gp120 preparation in cell growth medium, or supernatant from the stably-transfected cells, are incubated for 1 hour. plates are washed again, and incubated for one hour with a 10 human MoAb (F105, AIDS Research & Reference Reagent Program, No. 857). The plates are washed again, and incubated again for 1 hour with a horseradish-peroxidase-conjugated goat anti-human IgG (Cappel). Following a final wash, the peroxidase substrate OPD (DuPont) is added and the amount of qp120 determined by comparing absorbance of unknowns with a standard curve. Standards are prepared from purified gp120 made in CHO cells, a small quantity of which is obtained from Celltech Ltd. Clones expressing the highest levels are subjected to successive rounds of amplification of the newly introduced DNA sequences in increasing concentrations of 20 Stable CHO cell lines are thus generated methotrexate. which secrete at least 1 microgram/milliliter of mutant HIV-1 gp120.

25 11. Purification of HIV-1 qp120 proteins.

A one-step immunoaffinity procedure is used to purify the recombinant gp120 molecules described. Briefly, culture supernatant is collected and clarified by centrifugation. An immunoaffinity column consisting of a matrix coupled to a sheep polyclonal anti-gp120 IgG (D7234, Aalto Bioreagents) directed against the highly conserved C-terminal end (APTKAKRRVVQREKR) of gp120 is used to specifically adsorb gp120 from the cell culture media. This antisera recognizes native gp120, the V3 loop deletion mutants, and the CD4(4)

WO 94/22477 PCT/US94/03282

mutants since the C-terminal ends of these molecules remain The bound gp120 is then eluted with 2M MqCl. concentrated by Amicon filtration, and dialyzed into 10 mM HEPES, pH 7.0. The purity of the proteins is determined by 5 SDS-PAGE and silver staining.

43

Characterization of recombinant HIV-1 gp120 proteins. The purified glycoproteins are subjected to extensive biochemical and immunologic characterization. The integrity 10 of the proteins is monitored by SDS-PAGE and silver staining reducing and non-reducing conditions. glycoproteins are deglycosylated by treatment with the enzyme N-glycosidase F which cleaves N-linked oligosaccharides, and are assayed by SDS-PAGE and silver staining to monitor molecular weight shifts. The purified glycoproteins are also tested for reactivity with several well characterized anti-gp120 monoclonal antibodies that recognize both linear and discontinuous epitopes. binding affinity to sCD4 is estimated using an ELISA assay.

15

20

The purified proteins HIV-1 gp120_{LAI}, gp120_{LAI}-V3⁽⁻⁾, gp120_{LAI}-V3⁽⁻⁾ $^{\circ}$ -CD4 $^{\circ}$, gp120_{IR-FL}, gp120_{IR-FL}-V3 $^{\circ}$, and gp120_{IR-FL}-V3 $^{\circ}$ -CD4 $^{\circ}$, were tested for their ability to bind cell surface human CD4. DG44 #3 cells, a recombinant cell line designed to express 25 human CD4 on the membrane surface, were grown in T flasks and trypsinized. 5 X 105 cells/experiment were aliquoted into FACS buffer (PBS + 2% BSA and 0.1% NaN3), washed several times in the same buffer, and then incubated with 100 ul of a solution of purified gp120 protein at 5ug/ml in FACS buffer at 37°C for 2 hr. The cells were washed in FACS buffer, and then incubated in 100 ul solution containing 5ug/ml sheep polyclonal IgG against the highly conserved Cterminus of gp120 in FACS buffer at 37°C for 2 hr. cells were washed in FACS buffer then incubated in 100 ul

solution containing FITC-labeled rabbit anti-sheep IgG polyclonal antibody at 37°C for 2 hr. The cells were washed with FACS buffer and then resuspended in 500 ul FACS buffer. The cells were then analyzed on a Becton Dickinson FACScan according to the manufacturer's instructions. As a control for expression of CD4 on the DG44 #3 cells, FITC-labeled OKT4A (Becton Dickinson) was used.

13. A protocol for inoculation of animals with the mutant HIV-1 gp120 envelope glycoproteins.

Alum is used as an adjuvant during the inoculation series. The inoculum is prepared by dissolving the mutant HIV-1 gp120 envelope glycoprotein antigen in physiologic saline at a final antigen concentration of 100 ug/ml. Preformed alum (aluminum hydroxide gel) is added to the solution to a final level of 500 ug/ml aluminum. The antigen is allowed to adsorb onto the alum gel for two hours at room temperature. Following adsorption, the gel with the antigen is washed twice with physiologic saline and resuspended in the saline to a protein concentration of 100 ug/ml.

Monkeys and/or Guinea Pigs are individually inoculated with four 100 ug doses of the mutant HIV-1 gp120 envelope glycoprotein antigen adscrbed onto alum. Each dose is injected intramuscularly. The doses are delivered one or five months apart (week 0, 4, 8 and 28). the animals are bled at intervals of two or four weeks. Serum samples are prepared from each bleed to assay for the development of specific antibodies as described in the subsequent sections.

30

35

25

14. Analysis of sera for anti-mutant HIV-1 gp120 envelope glycoprotein IqG antibodies.

Each serum sample is analyzed by ELISA. Polystyrene microtiter plates are coated with 0.5 ug per well of pure mutant HIV-1 gp120 envelope glycoprotein in phosphate-

buffered physiological saline (PBS) at 4°C. Each well is then washed with PBS containing 0.5% TWEEN-20 (PBS-TW). Test serum, diluted serially in PBS-TW, is added to the mutant HIV-1 gp120 envelope glycoprotein-containing wells 5 and allowed to react with the adsorbed mutant HIV-1 gp120 envelope glycoprotein for one hour at 37°C. The wells are then washed extensively in PBS-TW. Each well then receives 0.1% p-nitrophenyl phosphate in 10% diethanolamine, pH 9.8 containing 0.5 mM MgCl₂.6H₂0. The ensuing reaction is allowed to proceed at room temperature for 30 minutes, at which time it is terminated by the addition of 3.0 N NaOH.

The greater the interaction of antibodies in the test serum with the mutant HIV-1 gp120 envelope glycoprotein, the 15 greater is the amount of alkaline phosphatase bound onto the The phosphatase enzyme mediates the breakdown of pnitrophenyl phosphate into a molecular substance which absorbs light at a wavelength of 405 nm. Hence, there exists a direct relationship between the absorbance at 405 20 nm of light at the end of the ELISA reaction and the amount of mutant HIV-1 gp120 envelope glycoprotein-bound antibody. All animals inoculated with mutant HIV-1 gp120 envelope glycoprotein whose serum reacts specifically with the mutant HIV-1 gp120 envelope glycoprotein in the ELISA have a 25 positive antibody response against mutant HIV-1 gp120 envelope glycoprotein.

Analysis of sera for activity which specifically neutralizes HIV-1 infectivity.

Virus-neutralizing activity is determined with an assay 30 based on the use of multiplicity curves in which the ratio of infectious virus surviving antibody treatment (V_n) is compared to infectious virus in uninhibited cultures (V.) at various dilutions of antisera. The neutralization titer of

the sera is then interpolated as that sera dilution which yields one log reduction in infectious titer (i.e., $V_n/V_o =$ Briefly, 4-fold dilutions of virus (laboratoryadapted and primary isolates) are prepared to yield 5 infectious doses of 0.1 to 100 TCID₅₀ (Tissue Culture Infection Dose) in 20 ul. Serial 3-fold dilutions of sera are also prepared and 20 ul of each serum dilution are incubated with each dilution of virus in duplicate for 60 minutes at room temperature in a 96-well microtiter plate. 20 ul of AA5 cells (PHA stimulated PBMCs for primary HIV-1 10 isolates) are then added to the serum/virus mixtures. Cells are cultured for 7 days by the addition of fresh medium every other day. On the seventh day, supernatant from each well is removed and tested for the presence of reverse 15 transcriptase (RT). Infection in each well is then scored as either positive or negative based on the RT counts, and the infectious dose of virus in each treatment group is calculated using the Reed and Muench (28) formula. neutralization titers represent the reciprocal serum 20 dilution required to reduced infectious dose of virus by one The above culture time is for the prototypic HIV-11A1 isolate tested on the AA5 cell line. In the case of primary isolates, the termination date is usually 11-14 days. Culture conditions for PBMCs is not as demanding since 25 doubling time is restricted. In the case of PBMCs, one day PHA stimulations are used at a final concentration of 1.5 X 106/ml on day 0. Half that number of fresh PBMCs are then added again on days 4 and 8. This multiple addition of PBMCs is meant to amplify virus output upon successful 30 infection so that the readout RT signal is strong. the final readout titer for the primary isolate/PBMC is the reciprocal serum dilution which reduces infectious titer by one log.

16. Passive hyperimmune therapy.

Non-HIV-1-infected humans are immunized with the mutant HIV-1 gp120 envelope glycoprotein antigens according to a protocol similar to that described above in section 12. For passive hyperimmune therapy in HIV-1-infected individuals, blood plasma is taken from mutant HIV-1 gp120 envelope glycoprotein immunized, non-HIV-1-infected human donors whose plasma has high levels of neutralizing antibodies. The plasma is pooled from several donors, purified to remove nonimmunoglobulin proteins and is then sterilized to kill any other viruses or pathogens. The treated plasma is then injected into individuals infected with HIV-1, with repeated injections every week, every two weeks, or every month.

WO 94/22477 PCT/US94/03282

48

Results

Eukaryotic expression vectors designed to express high levels of HIV-1_{LAI} gp120 and HIV-1_{JR-FL} gp120 were constructed. 5 The CMV MIE promoter/enhancer was used to drive the transcription of a gene fusion consisting of the human tPA signal sequence fused to mature gp120 (Figures 2 and 7). The complete sequence of the transcription unit from the Hinc II site of the CMV promoter/enhancer to the Not I site 10 just 3' from the stop codon in gp120 is shown in figure 3. This vector was used to transfect COSM5 cells in a transient assay. The transfected cells were labeled with 35S-cysteine and the media immunoprecipitated with a CD4-immunoglobulin-Protein A-Sepharose complex. The precipitated products were 15 analyzed using a reducing 10% SDS-PAGE autoradiography (Figure 4). A 120 kD band was detected when PPI4-tPA-gp120_{LAI} was used to transfect COS cells (lane 3). A band migrating with a slightly lower molecular mass was detected when PPI4-tPA-gp120 $_{\rm JR-FL}$ was used to transfect COS 20 cells (lane 4). No radiolabeled products were detected in the mock infected cells. Using a sheep polyclonal antibody directed against the highly conserved C-terminal end of HIV-1 gp120 in an ELISA assay, the level of expression of HIV-1 gp120 was determined to be 2350 ng/ml.

25

The PPI4-tPA-gp120_{LM} vector was then used to stably transfect the dhfr CHO cell line DXB11. Two days post-transfection, the cells were plated at low density in nucleoside-free medium. Eight days post-transfection, surviving clones were isolated and expanded. Individual primary transfectants were tested for gp120 expression using the ELISA method described in the methods section. Several primary CHO transfectants expressed significant quantities (10-120 ng/ml) of gp120 (Figure 5). Three of the highest

expressing clones were then subjected to increasing concentrations of methotrexate in order to amplify, tandem, the copy number of the dhfr and gp120 genes. Cell lines were established that express high levels of gp120 5 with rates of secretion greater than 1 mg/liter. These were then used to purify gp120 to homogeneity.

Six CHO cell lines were established, using the procedures described in the methods sections, that express high levels 10 of the following proteins: HIV-1 gp120_{LAI}, gp120_{LAI}-V3⁽⁻⁾, $gp120_{LAI}-V3^{(-)}-CD4^{(-)}$, $gp120_{IR-FL}$, $gp120_{IR-FL}-V3^{(-)}$, and $gp120_{IR-FL}-V3^{(-)}-$ CD4(-). Metabolic labeling of these cells with 35S-cysteine followed by immunoprecipitation with the human monoclonal antibody F105 and analyzed by SDS-PAGE and autoradiography showed the presence of the gp120 proteins in the culture supernatant (Figure 14). From these cell lines the gp120 proteins were purified to homogeneity. Analysis by SDS-PAGE followed by silver-staining showed the purity of these proteins to be greater than 90% (Figure 15).

20

25

15

It was shown by FACScan analysis that the two CD4 binding mutants HIV-1gp120_{LAI}-V3 $^{(\cdot)}$ -CD4 $^{(\cdot)}$ and HIV-1 gp120_{IR-FL}-V3 $^{(\cdot)}$ -CD4 $^{(\cdot)}$ had no appreciable binding to recombinant cell lines designed to express high levels of human CD4 on their membrane surface (Figure 16, panel 4 and data not shown, respectively)

Discussion

The advantage of using the mutant HIV-1 gp120 envelope glycoproteins as immunogens is that these proteins will not elicit an immune response against the V3 loop, a highly immunodominant epitope on gp120. This is significant because the V3 loop may skew the humoral immune response away from discontinuous epitopes in the CD4-binding site. Mutant HIV-1 10 gp120 envelope glycoproteins having partial and total v3 loop deletions have been made (30). Deletion of the V3 loop therefore exposes the CD4-binding site to the immune system, allowing the immune system to mount a response against this critical region (18). Another advantage of using the mutant 15 HIV-1 gp120 envelope glycoprotein as an immunogen is that it has significantly reduced affinity for cell surface CD4. An efficient humoral immune response depends on the binding of antigen to B cell surface immunoglobulin. The presence of the high-affinity CD4 receptor on large numbers of cells in the body may significantly diminish the ability of native gp120 to induce an effective humoral immune response. rationale of mutating gp120 at the CD4 binding site is to redirect the mutant HIV-1 gp120 envelope glycoprotein away from cell surface CD4 toward immunoglobulin-bearing B cells, thereby allowing the immune system to mount a response 25 against, inter alia, the CD4-binding site.

References

1. Klatzmann, D.R., et. al. (1990) Immunodeficiency Reviews 2, 43-66.

5

- 2. Lasky, L.A., et. al. (1987) Cell 50, 975-985.
- 3. Maddon, P.J., et. al. (1986) Cell 47, 333-348.
- 10 4. Maddon, P.J., et. al. (1988) Cell 54, 865-874.
 - 5. Maddon, P.J., et. al. (1985) Cell 42, 93-104,
- Maddon, P.J., et. al. (1987) Proc. Natl. Acad. Sci.
 U.S.A. 84, 9155-9159.
 - 7. Richardson, N.E., et. al. (1988) Proc. Natl. Acad. Sci. U.S.A. 85, 6102-6106.
- 20 8. Chao, B.H., et. al. (1989) J. Biol. Chem. 264, 5812-5817.
 - 9. Arthos, J., et. al. (1989) Cell 57, 469-481.
- 25 10. Wang, J., et. al. (1990) Nature 348, 411-418.
 - 11. Ryu, S.-E., et. al. (1990) Nature 348, 419-426.
- 12. Leonard, C.K., et. al. (1990) J. Biol. Chem. 265, 30 10373-10382.
 - Earl, P.L., et. al. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 648-652.
- 35 14. Helseth, E., et. al. (1991) J. Virol. 65, 2119-2123.

- 15. Bolognesi, D.P. (1990) TIBTech 8, 40-45.
- 16. Olshevsky, U., et. al. (1990) J. Virol. 64, 5701-5707.
- 5 17. Steimer, K.S., et. al. (1991) AIDS 5, S135-143.
 - 18. Wyatt, R., et. al. (1992) J. Virol. 66, 6997-7004.
- 19. Zolla-Pazner, S., et. al. (1992) Sem. in Virology 3,203-211.
 - 20. Steimer, K.S., et. al. (1991) Science 254, 105-108.
 - 21. Pollard, S.R., et. al. (1992) EMBO J. 11, 585-591.
- 22. Okayama, H. (1983) Mol. Cell. Biol. 3, 280-289.

20

25

- 23. Maniatis, T., et. al. (1990) Molecular Cloning, Vol. 1-3.
- 24. Thomsen, D.R., et. al. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 659-663.
 - 25. Pennica, D., et. al. (1983) Nature 301, 214-221.
- 26. Wain-Hobson, S., et. al. (1985) Cell 40, 9-17.
 - 27. Koyanagi, Y. (1987) Science 236, 819-822.
- 30 28. Reed, L.J. (1938) Am. J. Hyg., 27, 493-497.
 - 29. Cohn, E.J. et al., (1944) J. Clin. Invest. 23, 417-432.
- 30. Shiow-Her, C., et al. (1992) J. of Cellular Biochem.,

 Supplement 16E, Abstrtact Q105.

(1) GENERAL INFORMATION:

SEQUENCE LISTING

(i) APPLICANT: Progenics Pharmaceuticals, Inc. (ii) TITLE OF INVENTION: HIV-1 VACCINES, ANTIBODY COMPOSITIONS RELATED THERETO, AND THERAPEUTIC AND PROPHYLACTIC USES THEREOF (iii) NUMBER OF SEQUENCES: 29 (iv) CORRESPONDENCE ADDRESS: (A) ADDRESSEE: Cooper & Dunham (B) STREET: 30 Rockefeller Plaza (C) CITY: New York (D) STATE: New York (E) COUNTRY: USA (F) ZIP: 10112 (V) COMPUTER READABLE FORM: (A) MEDIUM TYPE: Floppy disk (B) COMPUTER: IBM PC compatible (C) OPERATING SYSTEM: PC-DOS/MS-DOS (D) SOFTWARE: Patentin Release #1.24 (vi) CURRENT APPLICATION DATA: (A) APPLICATION NUMBER: (B) FILING DATE: (C) CLASSIFICATION: (vii) PRIOR APPLICATION DATA: (A) APPLICATION NUMBER: US 08/037,816 (B) FILING DATE: 26-MAR-1993 (viii) ATTORNEY/AGENT INFORMATION: (A) NAME: White, John P. (B) REGISTRATION NUMBER: 28,678 (C) REFERENCE/DOCKET NUMBER: 41190-A-PCT/JPW/AJM (ix) TELECOMPAINICATION INFORMATION: (A) TELEPHONE: (212) 977-9550 (B) TELEFAX: (212) 664-0525 (C) TELEX: 422523 COOPUI (2) INFORMATION FOR SEQ ID NO:1: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 45 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1: X20 XAA XAB CYS XAB ILE XAB XAB XAB XAB XAB Trp XAB XAB XAB Xaa Xaa Ala Xaa Tyr Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Ser xaa xaa Thr Gly xaa xaa xaa xaa Arg xaa Gly xaa

35 40

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Thr Leu Pro Cys Arg Ile Lys Gln Phe Ile Asn Met Trp Gln Glu Val

Gly Lys Ala Met Tyr Ala Pro Pro Ile Ser Gly Gln Ile Arg Cys Ser 20 25 30

Ser Ash lie Thr Gty Leu Leu Leu Thr Arg Asp Gty Gty 35 40

- (2) INFORMATION FOR SEQ ID NO:3:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Thr Leu Pro Cys Arg Ile Lys Gin Ile Ile Asn Met Trp Gin Glu Val

Gly Lys Ala Met Tyr Ala Pro Pro Ile Arg Gly Gln Ile Arg Cys Ser 20 25 30

Ser Asn Ite Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly 35 40 45

- (2) INFORMATION FOR SEG ID NO:4:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 45 base pairs
 - (B) TYPE: nucleic scid (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATCCTGCAG TCACCGTCCT TGACACGATG GATGCAATGA AGAGA

- (2) INFORMATION FOR SEG ID NO:5:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid

(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:	
AAGTETTETE CTEGGTETTG TETTTTTAAC ACCEAG	36
(2) INFORMATION FOR SEQ ID NO:6:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 36 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:	
TTCAGAAGAG GAGCCAGAAC AGAAAAATTG TGGGTC	. 36
(2) INFORMATION FOR SEQ ID NO:7:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 39 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7: GGAAAAAAGC GGCCGCTCAT TTTTCTCTCT GCACCACTC	39
(2) INFORMATION FOR SEG ID NO:8:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 33 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	33
(2) INFORMATION FOR SEG ID NO:9:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 49 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	

(%1) ZEGNENCE DEZCKINITON: ZEG ID MO:A:	
CTGTAGAAAT TAATTGTACA GGTGCTGGAC ATTGTAACAT TAGTAGAGC	49
(2) INFORMATION FOR SEQ ID NO:10:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 29 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
CTCGAGCATG CATTCGAAGC TCGCTGATC	29
(2) INFORMATION FOR SEQ ID NO:11:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 29 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:	
CAATTTATAA ACATGGTGCA GGAAGTAGG	29
(2) INFORMATION FOR SEQ ID NO:12:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 29 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
CAAATTATAA ACATGGTGCA GGAAGTAGG	29
(2) INFORMATION FOR SEG ID NO:13:	
(i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3125 base pairs (B) TYPE: nucleic acid (C) STRANDEDMESS: single (D) TOPOLOGY: linear	
(ii) MOLECULE TYPE: DNA (genomic)	
(ix) FEATURE: (A) HAME/KEY: CDS (B) LOCATION: 15553115 (D) OTHER INFORMATION:	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TIGACATTGA TTATTGACTA GTTATTAATA GTAATCAATT ACGGGGTCAT TAGTTCA	TAG 60
CCCATATATG GAGTTCCGCG TTACATAACT TACGGTAAAT GGCCCGCCTG GCTGACC	:GCC 120
CAACGACCCC CGCCCATTGA CGTCAATAAT GACGTATGTT CCCATAGTAA CGCCAAT	AGG 180
GACTITICCAT TGACGICAAT GGGTGGACTA TTTACGGTAA ACTGCCCACT TGGCAGT	TACA 240
TCAAGTGTAT CATATGECAA GTACGEEECE TATTGAEGTE AATGAEGGTA AATGGE	CGC 300
CTGGCATTAT GCCCAGTACA TGACCTTATG GGACTTTCCT ACTTGGCAGT ACATCTA	ACGT 360
ATTAGTCATC GCTATTACCA TGGTGATGCG GTTTTGGCAG TACATCAATG GGCGTGG	ATA 420
GCGGTTTGAC TCACGGGGAT TTCCAAGTCT CCACCCCATT GACGTCAATG GGAGTTT	GTT 480
TTGGCACCAA AATCAACGGG ACTITCCAAA ATGTCGTAAC AACTCCGCCC CATTGAC	GCA 540
AATGGGCGGT AGGCGTGTAC GGTGGGAGGT CTATATAAGC AGAGCTCGTT TAGTGAA	ACCG 600
TCAGATCGCC TGGAGACGCC ATCCACGCTG TTTTGACCTC CATAGAAGAC ACCGGGA	ACCG 660
ATCCAGCCTC CGCGGCCGGG AACGGTGCAT TGGAACGCGG ATTCCCCGTG CCAAGAG	TGA 720
CGTAAGTACC GCCTATAGAC TCTATAGGCA CACCCCTTTG GCTCTTATGC ATGCTAT	ACT 780
GTTTTTGGCT TGGGCCAACA CCCCGTCCTA GATAGGTGAT GGTATAGCTT AGCCTAT	AGG 840
TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG ACGATACTTT CCATTAC	TAA 900
TCCATAACAT GGCCGCTCTT TGCCACAACT ATCTCTATTG GCTATATGCC AATACTC	TGT 960
CCTTCAGAGA CTGACACGGA CTCTGTATTT TTACAGGATG GGGTCCCATT TATTATT	TAC 1020
AMATTCACAT ATACAACAAC GCCGTCCCCC GTGCCCGCAG TTTTTATTAA CATGCGG	GAT 1080
CTCCACGCGA ATCTCGGGTA CGTGTTCCGG ACATGGGCTC TTCTCCGGTA GCGGCGG	AGC 1140
TECACATECG AGESTGTESS ATGSSSCATGS STEERAGEGGS TEATGGTEGS TEGGSCAS	iCTC 1200
CTTECTCCTA ACAGTGGAGG CCAGACTTAG GCACAGGACA ATGCCCACCA CCACCAG	TGT 1260
GCCGCACAAG GCCGTGGCGG TAGGGTATGT GTCTGAAAAT GAGCTCGGAG ATTGGGG	TCG 1320
CACCGCTGAC GCAGATGGAA GACTTAAGGC AGCGGCAGAA GAAGATGCAG GCAGCTG	AGT 1380
TGTTGTATTC TGTAGAGTTG GAGGTAACTC CCGTTGCGGT GCTGTTAACG GTGGAGG	GCA 1440
STGTAGTCTG AGCAGTACTC GTTGCTGCCG CGCGCGCCAC CAGACATAAT AGCTGAC	AGA 1500
CTAACAGACT GTTCCTTTCC ATGGGTCTTT TCTGCAGTCA CCGTCCTTGA CACG AT Me	
CAT GCA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG TGT GGA GO ASP Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly Al 5 10 15	
OTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA GC Val Phe Val Ser Pro Ser Gln Glu lle His Ala Arg Phe Arg Arg G 20 25 30	GC 1653 ly
GCC AGA ACA GAA AAA TTG TGG GTC ACA GTC TAT TAT GGG GTA CCT G Na Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Va 35 40 45	rg 1701 al
IGG AAG GAA GCA ACC ACC ACT CTA TTT TGT GCA JCA GAT GCT AAA GC	CA 1749

7 m		s Gl	u Ala	a Th	r Thi	r Thi	Le	. Phe	e Cys	Ala 60	Asp	Ala	Lys	Ala 65	
					l His	AAT S ASP				Thr					1797
				Pro		GAA Glu			Leu						1845
			Trp			GAC ASP		Val							1893
		Leu				AGC Ser 120									1941
	Cys					TGE Cys									1 989
						AGT Ser								AAA Lys	2037
						TCT Ser									2085
						GCA Ala					 				2133
						AGC Ser 200				Thr					2181
						CCA Pro									22 29
						GGT Gly									2277
						CCA Pro	Cys								2325
						GTA Val									2373
					Glu	GTA (Val 1 280				Ser					2421
				Ile		GTA (Val (Asn						2469
			Pro			MAT /		Arg							2517
						GTT /									2565

59 -

			325					330					335	,		٠	
GLD	GCA Ata	CAT His 340	TGT Cys	AAC Asn	ATT	AGT Ser	AGA Arg 345	GCA Ala	AAA Lys	TGG Trp	AAT Asn	GCC Ala 350	ACT Thr	TTA Leu	AAA Lys	;	2613
CAG Gln	ATA 1 (e 355	GCT Ala	AGC Ser	AAA Lys	TTA Leu	AGA Arg 360	GAA Gl·u	CAA Gln	TTT Phe	GGA Gly	AAT Asn 365	AAT Asn	AAA Lys	ACA Thr	ATA ile	;	2661
ATC Ile 370	TTT Phe	AAG Lys	CAA Gln	TCC Ser	TCA Ser 375	GGA Gly	GGG Gly	GAC ASP	CCA Pro	GAA Glu 380	ATT ile	GTA Val	ACG Thr	CAC His	AGT Ser 385	;	2709
TTT Phe	AAT Asn	TGT Cys	GGA Gly	GGG Gly 390	GAA Glu	TTT Phe	TTC Phe	TAC Tyr	TGT Cys 395	AAT Asn	TCA Ser	ACA Thr	CAA Gln	CTG Leu 400	TTT Phe	;	2757
AAT Asn	AGT Ser	ACT Thr	TGG Trp 405	TTT Phe	AAT Asn	AGT Ser	ACT Thr	1GG 1rp 410	AGT Ser	ACT	GAA Glu	GGG Gly	TCA Ser 415	AAT Asn	AAC Asn		2805
Thr	Glu	Gly 420	Ser	GAC ASP	Thr	lle	Thr 425	Leu	Pro	Cys	Arg	11e 430	Lys	Gln	Pne		2853
Ile	Asn 435	Met	Тгр	CAG Gln	Glu	Val 440	Gly	Lys	Ala	Met	Tyr 445	Ala	Pro	Pro	ite		2 9 01
Ser 450	Gly	Gln	lle	AGA Arg	Cys 455	Ser	Ser	Asn	Ile	Thr 460	Gly	Leu	Leu	Leu	1hr 465		2949
Arg	Asp	Gly	Gly	AAT Asn 470	Asn	Asn	Asn	Gly	Ser 475	Glu	Ile	Phe	Arg	Pro 480	Gly		2997
Gly	Gly	Asp	Met 485	AGG Arg	Asp	Asn	īгр	Arg 490	Ser	Glu	Leu	Туг	Lys 495	Туг	Lys		3045
Val	Val	Lys 500	Ite	Glu	Pro	Leu	Gly 505	Val	Ala	Pro	ACC Thr	AAG Lys 510	GCA	AAG Lys	AGA		3093
				AGA Arg			TG	\GCG(SCCG								3125

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 520 amino acids (B) TYPE: amino acid

 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 15 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60 Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 55 70 75 80 Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95 Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110 Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125 Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140 Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Met Glu
145 150 155 160 Lys Gly Glu Ile tys Asn Cys Ser Phe asn fie Ser Thr Ser fle arg Gly Lys Val Gin Lys Glu Tyr Ala Phe Phe Tyr. Lys Leu Asp Ile Ile 180 185 190 Pro 11e Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205 Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220 Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240 Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255 Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Asn 260 265 270 Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr 275 280 285 Asp Asn Ala Lys Thr Ile Ile Val Gin Leu Asn Gin Ser Val Giu Ile 290 295 300 Asn Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gln 305 310 315 Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Gly Lys Ile Gly Asn Met 325 330 335 Arg Gln Ala His Cys Asn Ile Ser Arg Ala Lys Trp Asn Ala Thr Leu 340 345 350 Lys Gin Ite Ala Ser Lys Leu Arg Glu Gin Phe Gly Asn Asn Lys Thr 355 360 365 lle lle Phe Lys Gln Ser Ser Gly Gly Asp Pro Glu lle Val Thr His 370 386 Ser Phe Asn Cys Gly Gly Glu Phe Phe Tyr Cys Asn Ser Thr Gln Leu 385 390 395 400

Phe Asn Ser Thr Trp Phe Asn Ser Thr Trp Ser Thr Glu Gly Ser Asn

61 .

				40	5				410)				415		
Asr	Thr	Glu	Gly 420		r Asp) Thr	·ile	1hr 425) Pro	Cys	Arg	11e 430	Lys	Gln	·
Phe	: Ile	435		Tr	Glr	n Glu	≀ Val 440		/ Lys	Ala	Met	Tyr 445	Ala	Pro	Pro	
ite	Ser 450	Gly	Gin	H	e Arg	Cys 455		Ser	ASF	ile	Thr 460		Leu	Leu	Leu	
Thr 465	-	Asp	Gly	Gly	470		Asn	Asn	Gly	Ser 475	Glu	ile	Phe	Arg	Pro 480	
Gly	Gly	Gly	Asp	Met 485		Asp	Asn	Trp	490	Ser	Glu	Leu	Tyr	L ys 495	Tyr	
Lys	Val	Val	Lys 500		Glu	Pro	Leu	61 y 505		Ala	Pro	Thr	Lys 510	Ala	Lys	
Arg	Arg	Val 515	Val	Gln	Arg	Glu	Lys 520								٠	
(2)	INF	DRMA'	TION	FO?	SEQ	ID I	NO:1	5:								
		(1	A) LI B) TY C) S1 O) T(ENGT YPE: TRANI OPOL	H: 1: NUC DEDNI DGY:	532 (leic ESS: lim	base Boid Sing	paid gle								
	•	(E	1) NA 3) LO 0) OT	ME/I CAT HER	ION: INF	11 DRMA1	TION:		. n. n.	n. 15.						
	• •	SEC														
		GCA Ala														48
		TTC Phe														96
GGC Gly	GGC Gly	AGA Arg 35	GTA Val	GAA Glu	AAG Lys	TTG L eu	TGG Trp 40	GTC Val	ACA Thr	GTC Val	TAT Tyr	TAT Tyr 45	GGG	GTA Val	CCT Pro	144
GTG Val	TGG Trp 50	Lys	GAA Glu	GCA Ala	ACC Thr	ACC Thr 55	ACT Thr	CTA Leu	TTT Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT ASP	GCT	AAA Lys	192
GCA Ala 65	TAT	GAT ASP	ACA Thr	GAG Glu	GTA Val 70	CAT His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	ACA Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	CCA Pro	CAA Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GAA Glu	AAT Asn	GTA Val	ACA Thr 95	GAA Glu	288
CAT	TTT	AAC	ATG	TGG	**	AAT	AAC	ATG	GTA	GAA	CAG	ATG	CAG	GAG	GAT	336

			Leu		GAT ASP			Leu								384
		Cys			TTA Leu		Cys									432
	ASI				GGA Gly 150											480
TCT Ser	TTC Phe	AAT Asn	ATC	ACC Thr 165	ACA Thr	AGC Ser	ATA Ile	AGA Arg	GAT Asp 170	GAG Glu	GTG Val	CAG Gln	AAA Lys	GAA Glu 175	TAT Tyr	528
GCT Ala	CTT Leu	TTT Phe	TAT Tyr 180	Lys	CTT Leu	GAT Asp	GTA Val	GTA Val 185	CCA Pro	ATA Ile	GAT Asp	AAT Asn	AAT Asn 190	AAT Asn	ACC Thr	576
					AGT Ser											624
					GAG Glu											672
					AAG Lys 230											720
					AGC Ser											768
					CTG Leu											816
_	-				GAC Asp											864
					TCT Ser											912
				Ser	ATA Ile 310											960
			He		GGA GLY											1008
AGA Arg	GCA Ala	AAA Lys	TGG Trp 340	AAT Asn	GAC Asp	ACT Thr	Leu	AAA Lys 345	CAG Gln	ATA Ile	GTT Val	ATA Ile	AAA Lys 350	Leu	AGA Arg	1056
GAA Glu	CAA Gln	111 Phe 255	GAG Glu	AAT Asn	AAA Lys	Thr	ATA Ile 360	GTC Val	TTT Phe	AAT Asn	CAC His	TCC Ser 365	TCA Ser	GGA	GGG	1104
GAC Asp	CCA Pro 370	GAA Glu	ATT Ile	GTA Val		CAC His 375	AGT Ser	TTT Phe	AAT Asn	TGT Cys	GGA Gly 380	GGA Gly	GAA Glu	TTT Phe	TTC Phe	1152
TAC	TGT	AAT	TCA	ACA	CAA	CTG	111	AAT	AGT	ACT	TGG	AAT	AAT	AAT	ACT	1200

1 yr 385		Asr	Ser	Thr	Gln 390		Phe	Asn	Ser	Thr 395		Asn	Asn	Asn	Thr 400	
,,,,					• • •										400	
GAA	GGG	TCA	AAT	AAC	ACT	GAA	GGA	AAT	ACT	ATC	ACA	CTC	CCA	TGC	AGA	1248
Glu	Gly	Ser	Asn	Asn		Glu	Gly	Asn			Thr	Leu	Pro	•	Arg	
				405					410					415		
ATA	444	CEA	417	ATA	226	ATG	166	č ko	CAA	GTA	GGA	462	CCA	475	TAT	1296
				Ite												, 270
	-,5	••••	420	• • •				425		•••	•.,	-,5	430		. ,,	
acc	CCT	ccc	ATE	AGA	GGA	CAA	ATT	A.C.A	TGT	TCA	TCA	AAT	ATT	474	ccc	1344
				Arg												
		435		3	,		440	3	-,-			445			,	
CTG	CTA	TTA	ACA	AGA	GAT	GGT	GGT	ATT	AAT	GAG	AAT	GGG	ACC	GAG	ATC	1392
Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	ile	Asn	Glu	Asn	Gly	Thr	Glu	lle	
	450					455					460					
TTC	AGA	CCT	GGA	GGA	GGA	GAT	ATG	AGG	GAC	AAT	TGG	AGA	AGT	GAA	TTA	1440
Phe	Arg	Pro	Gly	Gly		Asp	Het	Arg	ASP		Trp	Arg	Ser	GLU	Leu	
465					470					475					480	
TAT	AAA	TAT	AAA	GTA	GTA	AAA	ATT	GAA	CCA	TTA	GGA	GTA	GCA	acc	ACC	:488
Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile	Glu		Leu	Gly	۷a۱	Ala		Thr	
				485					490					495		
AAG	GCA	AAG	AGA	AGA	GTG	GTG	CAA	AGA	GAA	AAA	T GA	GCGC	CCGC	:		1532
Lys	Ala	Lys	Arg 500	Arg	Val	Vai	Gln	Arg 505	Glu	Lys						

- (2) INFORMATION FOR SEQ ID NO:16:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 507 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro $35 \hspace{1cm} 40 \hspace{1cm} 45 \hspace{1cm} .$

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Glm Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95

His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Gln Glu Asp 100 105 110

Ite Ite Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr Asn Asp Ser Glu Gly Thr Met Glu Arg Gly Glu Ile Lys Asn Cys 145 150 155 160 Ser Phe Asn I le Thr Thr Ser I le Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175 Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190 Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gln Ala Cys 195 200 205 Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala 210 215 220 Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly 225 230 235 240 Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly 1le Arg Pro 245 250 255 Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ata Glu Glu Glu 260 _ 265 270 Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile 1le 275 280 285 Val Gin Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300 Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr 305 310 315 320 Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser 325 330 335 Arg Ala Lys Trp Asn Asp Thr Leu Lys Gln Ile Val Ile Lys Leu Arg 340 345 350 Glu Gln Phe Glu Asn Lys Thr Ile Val Phe Asn His Ser Ser Gly Gly 355 360 365 Asp Pro Glu Ile Val Met His Ser Phe Asn Cys Gly Gly Glu Phe Phe 370 375 380 Tyr Cys Asn Ser Thr Gln Leu Phe Asn Ser Thr Trp Asn Asn Asn Thr 385 390 395 400 Glu Gly Ser Asn Asn Thr Glu Gly Asn Thr Ile Thr Leu Pro Cys Arg 405 410 415 Tie Lys Gin Tie Tie Asn Met Trp Gin Glu Val Gly Lys Ala Met Tyr 420 425 430 Ala Pro Pro 1le Arg Gly Gln 1le Arg Cys Ser Ser Asn 1le Thr Gly
435 440 445 Leu Leu Eur Thr Arg Asp Gly Gly 11e Asn Glu Asn Gly Thr Glu 11e 450 455 460 Phe Arg Pro Gly Gly Gly Amp Met Arg Amp Asn Trp Arg Ser Glu Leu 465 470 475 480 Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr 485 490 495 Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys 500 505

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1484 base pairs
 - (B) TYPE: nucleic acid (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (11) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..1474
 (D) OTHER INFORMATION:

.... CECHENCE DESCRIPTION. CEO ID NO.17

(xi) SE	QUE	ICE [ESCR	IPTI	ON:	SEO	ID N	0: 17	:				
Asp				Arg	GGG				Val				GGA Gly	48
 			;er		AGC Ser									96
		Thr			TIG Leu									144
	Lys				ACC Thr 55									192
					CAT									240
					CAA Gln									288
					AAT Asn									336
					CAA Gln									384
					AAG Lys 135									432
 					AAT Asn						-	 		480
					TGC Cys									528
					TAT Tyr									576
					ACC Thr									624

. •

TCA Ser	GTC Val 210	He	ACA Thr	CAC Gir	GCC Ala	TGT Cys 215	Pro	AAG Lys	GTA Val	TCC Ser	TTT Phe 220	GAG Glu	CCA Pro	ATT !le	CCC Pro	672	
ATA Ile 225	CAT	TAT Tyr	TGT Cys	GCC	CCG Pro 230	Ala	GGT Gly	TTT Phe	GCG Ala	ATT Ile 235	CTA Leu	AAA Lys	TGT Cys	AAT Asn	AAT Asn 240	720	
AAG Lys	ACG Thr	TTC Phe	AAT Asn	GGA Gly 245	ACA Thr	GGA Gly	CCA Pro	TGT Cys	ACA Thr 250	AAT Asn	GTC Ve l	AGC Ser	ACA Thr	GTA Val 255	CAA Gln	768	
TGT Cys	ACA Thr	CAT	GGA Gly 260	Ile	AGG Arg	CCA Pro	GTA Val	GTA Val 265	TCA Ser	ACT Thr	CAA Gln	CTG Leu	CTG Leu 270	TTG Leu	AAT Asn	816	
GGC	AGT Ser	CTA Leu 275	GCA Ala	GAA Glu	GAA Glu	GAG Glu	GTA Val 280	GTA Val	ATT	AGA [.] Arg	TCT Ser	GCC Ala 285	AAT Asn	TTC Phe	ACA Thr	864	
GAC Asp	AAT Asn 290	GCT Ala	AAA Lys	ACC Thr	ATA	ATA Ile 295	GTA Val	CAG Gln	CTG Leu	AAC Asn	CAA Gln 300	TCT	GTA Val	GAA Glu	ATT Ile	912	
AAT Asn 305	TGT Cys	ACA Thr	GGT Gly	GCT Ala	GGA Gly 310	CAT	TGT Cys	AAC Asn	ATT Ile	AGT Ser 315	AGA Arg	GCA Ala	AAA Lys	TGG Trp	AAT Asn 320	960	
					ATA Ile											1008	
					TTT Phe											1056	
					AAT Asn											1104	
Thr					AGT Ser											1152	
GGG Gly 385	TCA Ser	AAT Asn	AAC Asn	ACT Thr	GAA Glu 390	GGA Gly	AGT Ser	GAC Asp	ACA Thr	ATC Ite 395	ACA Thr	CTC Leu	CCA Pro	TGC Cys	AGA Arg 400	1200	
ATA Ile	AAA Lys	CAA Gln	TTT Phe	ATA Ile 405	AAC Asn	ATG Het	TGG Trp	CAG Gln	GAA Glu 410	GTA Val	GGA Gly	Lys	GCA	ATG Met 415	TAT	1248	
Ala	Pro	Pro	1 le 420	Ser		Gln	Ile	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gly	1296	
CTG Leu	CTA Leu	TTA Leu 435	ACA Thr	AGA Arg	GAT Asp	GGT Gly	GGT Gly 440	AAT Asn	AAC Asn	AAC Asn	AAT Asn	GGG Gly 445	Ser	GAG	Ile	1344	
TTC Phe	AGA Arg 450	CCT Pro	GGA Gly	GGA Gly	Gly	GAT ASP 455	ATG Het	AGG Arg	GAC ASP	AAT Asn	TGG Trp 460	Arg	AGT Ser	GAA	TTA	1392	
TAT Tyr 465	AAA Lys	TAT Tyr	AAA Lys	Val	GTA Val 470	AAA Lys	ATT	GAA Glu	CCA Pro	TTA Leu 475	GGA Gly	GTA Val	GCA	Pro	ACC Thr 480	1440	
AAG	GCA	AAG	AGA	AGA	GTG	GTG	CAG	AGA	GAA	**	T G	AGCG	GCCG	C		1484	,

Lys Ala Lys Arg Arg Val Val Gin Arg Glu Lys 485 490

- (2) INFORMATION FOR SEQ ID NO:18:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 491 amino acids
 - (B) TYPE: amino acid (D) TOPOLOGY: linear

 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEO ID NO:18:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys . 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Het Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Tie Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 . 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu 145 150 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gin Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asp Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gin Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

Ile His Tyr Cys Ala Pro Ala Gly Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn 260 265 270

Gly Ser Leu Ala Glu Glu Glu Val Val Ile Arg Ser Ala Asn Phe Thr 275 280 285

ASP		Ala	Lys	Thr	lle		Val	Gln	Leu	Asn	Gln 300	Ser	val	Glu	lle	
Asn	290 Cys	Thr	Gly	Ala	Gly	295 His	Cys	Asn	lle	Ser		Ala	Lys	Trp	Asn	
305					310					315					320	
Ala	Thr	Leu	Lys	Gln 325	Ile	Ala	Ser	Lys	Leu 330	Arg	Glu	Gln	Phe	Gly 335	Asn	
Asn	Lys	Thr	1 l e 340	Į l e	Phe	Lys	Gln	Ser 345	Ser	Gly	Gly	Asp	Pro 350	Glu	lle	
Val	Thr	His 355	Ser	Phe	Asn	Cys	Gly 360	Gly	Glu	Phe	Phe	1 yr 365	Cys	Asn	Ser	
Thr	Gln 370	Leu	Phe	Asn	Ser	Thr 375	Trp	Phe	Asn	Ser	1hr 380	Trp	Ser	Thr	Glu	
Gly 385	Ser	Asn	Asn	Thr	Glu 390	Gly	Ser	Asp	Thr	1 l e 395	Thr	Leu	Pro	Cys	Arg 400	
Ile	Lys	Gln	Phe	11 e 405	Asn	Met	Trp	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415	Tyr	
Ala	Pro	Pro	ile 420	Ser	Gly	Gln	lie	Arg 425	Cys	Ser	Ser	Asn	1 l e 430	Thr	Gly	
Leu	Leu	L eu 435	Thr	Arg,	ASP	Gly	Gly 440	Asn	Asn	Asn	Asn	Gly 445	Ser	Glu	Ile	
Phe	Arg 450	Pro	Gly	Gly	Gly	Asp 455	Met	Arg	Asp	Asn	1rp 460	Arg	Ser	Glu	Leu	
1yr 465	Lys	Tyr	Lys	Val	Val 470	Lys	He	Glu	Pro	Leu 475	Gly	Val	Ala	Pro	Thr 480	
Lys	Ala	Lys	Arg	Arg 485	Val	Val	Gln	Arg	6lu 490	Lys						
(2)	INFO	ORMAT	ION	FOR	SEQ	ID I	10:19	?:								
	(i)	() E)	OUENC A) LE B) TY C) S1 D) TC	NGTH PE: RAND	nucl	48 leic \$\$:	acio sing	pai:	rs							
	(ii)) MOI	.ECUL	E TY	PE:	DNA	(ger	nomi	c)							
	(ix)	()	ATURE A) NA B) LO D) O1	ME/K	ON:	1		:							•	
	(xí) SE	OUENC	E DE	SCRI	PTIC) 	SEQ :	ID N	D: 19	:					
ATG Met 1	GAT Asp	GCA Ala	ATG Met	AAG Lys 5	AGA Arg	GGG Gly	CTC Leu	TGC Cys	TGT Cys 10	val	CTG Leu	CTG Leu	CTG Lev	TG1 Cy:	GGA Gly	48
GCA Ala	GTC Val	TTC Phe	GTT Val 20	TCG Ser	CCC Pro	AGC Ser	CAG Gln	GAA Glu 25	ATC	CAT His	GCC	CGA	Phe	AG	A AGA B Arg	96
GGC Gly	GGC Gly	AGA Arg 35	GTA Val	GAA Glu	AAG Lys	TTG Leu	TGG Trp 40	GTC Val	ACA Thr	GTC Val	TAT	TA1	r Gli	G GT	A CCT L Pro	144

GTG Vat	TGG Trp 50	AAA Lys	GAA Glu	GCA	ACC Thr	ACC Thr 55	ACT Thr	CTA Leu	TTT Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT Asp	GCT Ala	AAA Lys ·	192
GCA Ala 65	TAT Tyr	GAT Asp	ACA Thr	GAG Glu	GTA Val 70	CAT	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	ACA Thr	CAT His	GCC Ala	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC Pro	AAC Asn 85	CCA Pro	CAA Gln	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GAA Glu	AAT Asn	GTA Val	ACA Thr 95	GAA Glu	288
CAT His	ITT Phe	AAC Asn	ATG Met 100	TGG Trp	AAA Lys	AAT Asn	AAC Asn	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAG Gln 110	GAG Glu	GAT Asp	336
ATA	ATC 1le	AGT Ser 115	TTA Leu	TGG Trp	GAT Asp	CAA Gln	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC Thr	384
CCA Pro	CTC Leu 130	TGT Cys	GTT Val	ACT Thr	TTA Leu	AAT Asn 135	TGC Cys	AAG Lys	GAT Asp	GTG Val	AAT Asn 140	GCT Ala	ACT Thr	AAT Asn	ACC Thr	432
ACT Thr 145	AAT Asn	GAT ASP	AGC Ser	GAG Glu	GGA Gly 150	ACG Thr	ATG Met	GAG Glu	AGA Arg	GGA Gly 155	GAA Glu	ATA Ile	AAA Lys	AAC Asn	TGC Cys 160	480
TCT Ser	TTC Phe	AAT Asn	ATC 1le	ACC Thr 165	ACA Thr	AGC Ser	ATA Ile	AGA Arg	GAT Asp 170	GAG Glu	GTG Val	CAG Gln	AAA Lys	GAA Glu 175	TAT	528
GCT Ala	CTT Leu	TTT Phe	TAT Tyr 180	AAA Lys	CTT Leu	GAT Asp	GTA Val	GTA Vai 185	CCA Pro	ATA Ile	GAT Asp	AAT Asn	AAT Asn 190	AAT Asn	ACC Thr	576
AGC Ser	TAT Tyr	AGG Arg 195	TTG Leu	ATA Ile	AGT Ser	TGT Cys	GAC Asp 200	ACC Thr	TCA Ser	GTC Val	ATT Ile	ACA Thr 205	CAG Gln	GCC Ala	TGT Cys	624
CCA Pro	AAG Lys 210	ATA 1le	TCC Ser	TTT Phe	GAG Glu	CCA Pro 215	ATT	CCC Pro	ATA 1le	CAT His	TAT Tyr 220	TGT Cys	GCC Ala	CCG Pro	GCT Ala	672
GGT Gly 225	TTT Phe	GCG Ala	ATT Ile	CTA Leu	AAG Lys 230	TGT Eys	AAT Asn	GAT ASP	AAG Lys	ACG Thr 235	TTC Phe	AAT Asn	GGA Gly	AAA Lys	GGA Gly 240	720
CCA Pro	TGT Cys	AAA Lys	AAT Asn	GTC Val 245	AGC Ser	ACA Thr	GTA Val	CAA Gln	TGT Cys 250	ACA Thr	CAT His	GGA Gly	ATT	AGG Arg 255	Pro	768
GTA Val	GTA Val	TCA Ser	ACT Thr 260	CAA Gin	CTG Leu	CTG L e u	CTA Leu	AAT ABD 265	GLY	AGT Ser	CTA Leu	GEA Ala	GAA Glu 270	GLU	GAG Glu	816
GTA Val	GTA Val	ATT Ile 275	AGA Arg	TCT Ser	GAC Asp	AAT Asn	TTC Phe 280	ACG Thr	AAC Asn	AAT Asn	GCT Ala	Lys 285	Thr	ATA	ATA : Ile	864
GTA Val	CAG Gln 290	CTG Leu	AAA Lys	GAA Glu	TCT Ser	GTA Val 295	GAA Glu	ATT	AAT Asn	TGT Cys	ACA Thr 300	Gly	GCT Ala	GG/ Gly	CAT His	912
TGT Cys 305	AAC Asn	ATT Ile	AGT Ser	AGA Arg	GCA Ala 310	AAA Lys	TGG Trp	AAT Asn	GAC Asp	ACT Thr 315	Leu	Lys	CAC Glr	ATA	Val 320	960
ATA	***	ATT	AGA	GAA	CAA	TTT	GAG	AAT	**	ACA	ATA	GTC	: 111	ı AA'	r CAC	1008

1 l e	Lys	Leu	Arg	Glu 325	Gln	Phe	Glu	Asn	L ys 330	Thr	!le	Val	Phe	Asn 335	His	
TCC Ser	TCA Ser	GGA Gly	GGG Gly 340	GAC Asp	CCA Pro	GAA Glu	ATT	GTA Val 345	ATG Met	CAC His	AGT Ser	TTT Phe	AAT Asn 350	TGT Cys	GGA Gly	1056
GGA Gly	GAA Glu	717 Phe 355	TTC Phe	TAC Tyr	TGT Cys	AAT Asn	TCA Ser 360	ACA Thr	CAA Gln	CTG Leu	1771 Phe	AAT ASD 365	AGT Ser	ACT Thr	TGG Trp	1104
Asn	AAT Asn 370	AAT Asn	ACT Thr	GAA Glu	GGG	TCA Ser 375	AAT Asn	AAC Asn	ACT Thr	GAA Glu	GGA Gly 380	AAT Asn	ACT Thr	ATC	ACA Thr	1152
CTC Leu 385	CCA Pro	TGC Cys	AGA Arg	ATA Ile	AAA Lys 390	CAA Gln	ATT 1le	ATA Ile	AAC Asn	ATG Met 395	TGG Trp	CAG Gln	GAA Glu	GTA Val	GGA Gly 400	1200
AAA Lys	GCA Ala	ATG Met	TAT Tyr	GCC Ala 405	CCT Pro	CCE Pro	ATC Ile	AGA Arg	GGA Gly 410	CAA Gln	ATT	AGA Arg	TGT Cys	TCA Ser 415	TCA Ser	1248
AAT Asn	ATT Ile	ACA Thr	GGG Gly 420	CTG Leu	CTA Leu	TTA Leu	ACA Thr	AGA Arg 425	GAT Asp	GGT Gly	GGT Gly	áîT Ile	AAT Asn 430	GAG Glu	AAT Asn	12 9 6
Gly	ACC Thr	GAG Glu 435	ATC Ile	TTC Phe	AGA Arg	CCT Pro	GGA Gly 440	Gly	GGA Gly	GAT ASP	ATG Met	AGG Arg 445	GAC ASP	AAT ASD	TGG Trp	1344
AGA Arg	AGT Ser 450	GAA Glu	TTA Leu	TAT Tyr	AAA Lys	TAT Tyr 455	AAA Lys	GTA Val	GTA Val	AAA Lys	ATT Ile 460	GAA Glu	CCA Pro	TTA Leu	GGA Gly	13 9 2
GTA Val 465	GCA Ala	CCC Pro	ACC Thr	AAG Lys	GCA Ala 470	AAG L <u>y</u> s	AGA Arg	AGA Arg	GTG Val	GTG Val 475	CAA Gln	AGA Arg	GAA Glu	AAA Lys	TG	1439
AGC	GCC	GC														1448

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 479 amino acids
 - (8) TYPE: amino acid
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
- Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
- Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30
- Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45
- Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
 50 55 60
- Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80
- Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Glu Asn Val Thr Glu

				85					90					95	
His	Phe	Asn	Met 100		Lys	Asn	Asn	Met 105	Val	Glu	Gln	Net	Gln 110	Glu	Asp
He	ile	Ser 115	ren	Тгр	Asp	Gln	Ser 120	Leu	Lys	Pro	Cys	Val 125	Lys	Leu	Thr
Pro	Leu 130	Cys	Val	Thr	Leu	Asn 135	Cys	Lys	ASP	Val	Asn 140	Ala	Thr	Asn	Thr
Thr 145	Asn	Asp	Ser	Glu	Gly 150	Thr	Met	Glu	Arg	Gly 155	Glu	lle	Lys	Asn	Cys 160
			lle	165					170					175	
			197 180					185					190		
	-	195	Leu				200					205			
	210		Ser			215					220				
225			!le		230					235					240
	-	·	Asn	245					250					255	
			1hr 260					265					270		
		275	Arg		·		280					285			
	290		Lys			295					300				
305	Asn	Ile	Ser	Arg	310	Lys	Тгр	ASN	ASP	315	Leu	Lys	GLN	He	320
lle	Lys	Leu	Arg	Glu 325	Gln	Phe	Glu	Asn	330 330	Thr	Ile	Val	Phe	Asn 335	His
			Gly 340					345					350		
Gly	Glu	355	Phe				360					365			
Asn	370		Thr			375					380				
385			Arg		390					395					400
				405					410					415	
			420					425					430		Asn
Gly	Thr	Glu 435	lle	Phe	Arg	Pro	Gly 440	Gly	Gty	Asp	Met	Arg 445	Asp	Asn	Trp
	c	c		Tve	1 ve	Twe	i ve	Val	Val	l vs	11e	Glu	Pro	Leu	Gly

	450					455					460					
val 465	Ala	Pro	Thr	Lys	Ala 470		Arg	Arg	Val	Val 475	Gln	Arg	Glu	Lys		
(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:2	1:								
		() () ()	A) L B) T C) S D) T	YPE: TRAN OPOL	H: 1 DEDNI OGY:	484 leic ESS: lin	base acid sing ear	paid d gle								
					YPE:	DNA	(gei	nomic	=)							
	(ix	CI	A) N. B) L	E: AME/I OCAT THER	1 ON :	1		:								
	(xi)) SE	QUE N	CE DI	ESCR	PTIC	ON: 5	SEQ I	D NO	:21:	:					
ATG Met . 1	GAT ASP	GCA Ala	ATG Met	AAG Lys 5	AGA Arg	GGG Gly	CTC Leu	TGC Cys	TGT Cys 10	GTG Val	CTG Leu	C1G Leu	CTG Leu	7GT Cys 15	GGA Gly	48
GCA Ala	GTC Val	TTC Phe	GTT Val 20	Ser	CCC Pro	AGC Ser	CAG Gln	GAA Glu 25	ATC 1le	CAT His	GCC Ala	CGA Arg	TTC Phe 30	AGA Arg	AGA Arg	96
GGC Gly	GCC	AGA Arg 35	ACA Thr	GAA Glu	AAA Lys	TTG Leu	TGG Trp 40	GTC Val	ACA Thr	GTC Val	TAT Tyr	TAT Tyr 45	GGG Gly	GTA Val	CCT Pro	144
GTG Val	TGG Trp 50	AAG Lys	GAA Glu	GCA	ACC Thr	ACC Thr 55	ACT Thr	CTA Leu	TTT Phe	TGT Cys	GCA Ala 60	TCA Ser	GAT Asp	GCT Ala	AAA Lys	192
GCA Ala 65	TAT Tyr	GAT ASP	ACA Thr	GAG Glu	GTA Val 70	CAT His	AAT Asn	GTT Val	TGG Trp	GCC Ala 75	ACA Thr	CAT	GCC Ala	TGT Cys	GTA Val 80	240
CCC Pro	ACA Thr	GAC Asp	CCC	AAC Asn 85	CCA Pro	CAA Gin	GAA Glu	GTA Val	GTA Val 90	TTG Leu	GTA Val	AAT Asn	GTG Val	ACA Thr 95	GAA Glu	288
AAT Asn	TTT Phe	AAC A sn	ATG Met 100	TGG Trp	AAA Lys	AAT Asn	GAC Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	GAG Glu	GAT Asp	336
ATA Ile	ATC Ile	AGT Ser 115	TTA Leu	TGG Trp	GAT Asp	CAA Gln	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA Leu	ACC	384
CCA Pro	CTC Leu 130	TGT Cys	GTT Val	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT Asp	TTG Leu	GGG Gly 140	Asn	GCT	ACT Thr	AAT Asn	432
ACC Thr 145	AAT Asn	AGT Ser	AGT Ser	AAT Asn	ACC Thr 150	AAT Asn	AGT Ser	AGT Ser	AGC Ser	666 61 y 155	GAA Glu	ATG	ATG Met	ATG Met	GAG Glu 160	480
AAA Lys	GGA Gly	GAG Glu	ATA Ile	AAA Lys 165	AAC Asn	TGC Cys	TCT Ser	TTC Phe	AAT Asn 170	ATC	AGC Ser	ACA Thr	AGC Ser	ATA Ile 175	AGA	528
GGT Gly	AAG Lys	GTG Val	CAG Gin	AAA Lys	GAA Glu	TAT Tyr	GCA Ala	TTT Phe	TTT Phe	TAT Tyr	AAA Lys	CTT Leu	GAT ASP	ATA	ATA : Ile	576

73

		180	0				185	•		190			
		ASI			ACC Thr		Tyr						624
	. 116				7G7 Cys 215	Pro							. 672
His					GCT Ala								720
				Thr	GGA Gly								768
			Ile		Pro								816
		Ala			GAG Glu								854
	Ala				ATA 1 l e 295								912
					CAT His								960
					GCT Ala								1008
					AAG Lys								1056
					TGT Cys								1104
					ACT Thr 375								1152
					GGA Gly								1200
		Phe			ATG Met							Tyr	1248
 					CAA Gln						Thr		1296
Leu										Ser		ATC Ile	1344
				Gly								TTA Leu	1392

TAT AAA TAT AAA GTA GTA AAA ATT GAA CCA TTA GGA GTA GCA CCC ACC
Tyr Lys Tyr Lys Val Val Lys Ile Glu Pro Leu Gly Val Ala Pro Thr
465 470 475 480

AAG GCA AAG AGA AG AGTGGTGCAG AGAGAAAAAT GAGCGGCCGC 1484

(2) INFORMATION FOR SEQ ID NO:22:

Lys Ala Lys Arg

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 484 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

. Pro Thr Asp Pro Asn Pro Gin Glu Val Val Leu Val Asn Val Thr Glu 85 90 95

Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110

Ile Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu
145 150 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Vai Gin Lys Giu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr 195 200 205

Ser Val Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Glu Pro Ile Pro 210 215 220

lle His Tyr Cys Ala Pro Ala Ety Phe Ala Ile Leu Lys Cys Asn Asn 225 230 235 240

Lys Thr Phe Asn Gly Thr Gly Pro Cys Thr Asn Val Ser Thr Val Gln 245 250 255

Cys Thr His Gly Ile Arg Pro Val Val Ser Thr Gln Leu Leu Leu Asn

			260	l				265					270		
Gly	Ser	Leu 275	Ala	Glu	Glu	Glu	Va (280		Ile	Arg	Ser	A.la 285	Asn	Phe	Thi
Asp	Asn 290		Lys	Thr	Ile	i l e 295	val	Gln	Leu	Asn	Gln 300	Ser	Val	Glu	Ιle
Asn 305	Cys	Thr	Gly	Ala	Gly 310		Cys	Asn	1 l e	Ser 315	Arg	Ala	Lys	īrp	Asr 320
Ala	Thr	Leu	L ys	Gln 325	Ile	Ala	Ser	Lys	1 eu 330	Arg	Glu	Gln	Phe	Gly 335	Asr
Asn	Lys	Thr	I le 340	lle	Phe	Lys	Gln	Ser 345	Ser	Gly	Gly	Asp	Pro 350	Glu	H
val	Thr	нis 355	Ser	Phe	Asn	Cys	Gly 360	Gly	Glu	Phe	Phe	Туг 365	Cys	Asn	Ser
Thr	Gln 370	Leu	Phe	Asn	Ser	1hr 375	Trp	Phe	Asn	Ser	Thr 380	Trp	Ser	Thr	Gli
Gly 385	Ser	Asn	Asn	Thr	Glu 390	Gly	Ser	Asp	Thr	1 i e 395	Thr	Leu	Pro	Cys	400
Ile	Lys	Gln	Phe	1 l e 405	Asn	Met	Val	Gln	Glu 410	Val	Gly	Lys	Ala	Met 415	Tyı
Ala	Pro	Pro	1 le 420	Ser	Gly	Gin	Ile	Arg 425	Cys	Ser	Ser	Asn	11e 430	Thr	Gli
Leu	Leu	Leu 435	Thr	Arg	Asp	Gly	Gly 440		Asn	Asn	Asn	Gly 445	Ser	Glu	Ite
Phe	Arg 450	Pro	Gly	Gly		Asp 455	Het	Arg	Asp	Asn	Тг р 460	Arg	Ser	Glu	Le
1yr 465	Lys	Туг	Lys	Val	Val 470	Lys	lie	Glu	Pro	Leu 475	Gly	Val	Ala	Pro	Th:
Lys	Ala	Lys	Arg												
121	INSC	DMAT	TON	FOR	SED	ID N	n - 23								

- - (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1448 base pairs

 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: Linear
 - (ii) MOLECULE TYPE: DNA (genomic)
 - (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1..1438 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

- ATG GAT GGA ATG AAG AGA GGG CTC TGC TGT GTG CTG CTG CTG TGT GGA
 Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly
 1 5 10 15
- GCA GTC TTC GTT TCG CCC AGC CAG GAA ATC CAT GCC CGA TTC AGA AGA Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25

			Val			: TTG : Leu		Val								144
		Lys				ACC Thr 55	Thr									192
	Tyr					CAT										240
						CAA Gln										288
				Trp		AAT Asn										336
						CAA Gln										384
						AAT Asn 135										432
						ACG Thr										480
						AGC Ser										528
						GAT Asp										576
						TGT Cys										624
						CCA Pro 215										672
				Leu		TGT Cys									GGA Gly 240	720
			Asn												CCA Pro	768
GTA Val	GTA Val	Ser	ACT Thr 260	CAA Gln	CTG Leu	CTG Leu	CTA Leu	AAT Aan 265	GGC Gly	AGT Ser	CTA Leu	GCA Ala	GAA Glu 270	Glu	GAG	816
	Val					Asn							Thr		ATA	864
					Ser							Gly			CAT His	912
TGT	AAC	ATT	AGT .	AGA :	GCA	***	TGG	AAT	GAC	ACT	TTA	**	CAG	ATA	GTT	960

rve	Acn	tie	Ser	Aro	Ala	LVS	Tro	Asn	ASD	Thr	l eu	1 VS	Gin	He	Val	
305		.,.	30.	3	310			A3.,	, , , , , , , , , , , , , , , , , , ,	315		-,5	••••	•••	320	
			AGA													1008
Ite	Lys	L eu	Arg	Glu 325	Gln	Phe	Glu	Asn	Lys 330	Thr	lle	Val	Phe	Asn 335	His	
			GGG													1056
Ser	Ser	Gly	Gly 340		Pro	Glu	ile	Vál 345	Met	His	Ser	Phe	Asn 350	Cys	Gly	
GGA	GAA	TTT	TTC	TAC	TGT	AAT	TCA	ACA	CAA	CTG	TTT	AAT	AGT	ACT	TGG	1104
Gly	Glu	Phe 355	Phe	Tyr	Cys	Asn	Ser 360	Thr	Gln	Leu	Phe	Asn 365	Ser	Thr	Trp	
AAT	AAT	AAT	ACT	GAA	GGG	TCA	AAT	AAC	ACT	GAA	GGA	AAT	ACT	ATC	ACA	1152
Asn	Asn 370	Asn	Thr	Glu	Gly	Ser 375	Asn	Asn	Thr	Glu	Gly 380	Asn	Thr	He	Thr	
			AGA													1200
Leu 385	Pro	Cys	Arg	Ile	198 390	Gln	lie	Ile	Asn	Met 395	Val	Gln	Glu	Val	Gly 400	
			TAT													1248
Lys	Ala	Met	Туг	Ala 405		Pro	Ile	Arg	Gly 410	Gln	lle	Arg	Cys	Ser 415	Ser	
			GGG													1296
Asn	He	Thr	Gly 420	Leu	Leu	Leu	Thr	Arg 425	Asp	Gly	Gly	ile	Asn 430	Glu	Asn	
GGG	ACC	GAG	ATC	TTC	AGA	CCT	GGA	GGA	GGA	GAT	ATG	AGG	GAC	AAT	TGG	1344
Gly	Thr	Glu 435	Ite	Phe	Arg	Pro	Gly 440	Gly	Gly	Asp	Het	Arg 445	Asp	Asn	Trp	
			ATT						_							1392
Arg	Ser 450	Glu	l eu	Tyr		1yr 455	Lys	Val	Val	Lys	11e 460	Glu	Pro	Leu	Gly	
GTA	GCA	CCC	ACC	AAG	GCA	AAG	AGA	AGA	GTG	GTG	CAA	AGA	GAA	*	T	1438
-	Ala	Pro	Thr			Lys	Arg	Arg	Val		Gln	Arg	era	Lys	•	
465					470					475						
GAGC	GGCC	GC														1448

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 479 amino acids
 (B) TYPE: amino acid

 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Asp Ala Met Lya Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro
35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys
50 55 60

Ala 65		ASC	Thi	r Glu	yal 70		Asn	Val	Trp	Ata 75	Thr	His	Ala	Cys	Val 80
Pro	Thr	Asp	Pro	Asr 85		Gln	Glu	Vat	Val 90	Leu	Glu	Asn	Val	Thr 95	Glu
His	Phe	Asn	100	: Trp	lys	Asn	Asn	Met 105	Val	Glu	Gln	Met	Gln 110	Glu	Asp
Ile	Ile	Ser 115		, Trp	ASP	Gln	Ser 120	L eu	Lys	Pro	Cys	Val 125	Lys	Leu	Thr
Pro	L eu 130		Val	Thr	l eu	Asn 135	Cys	Lys	Asp	Val	Asn 140	Ala	Thr	Asn	Thr
Thr 145	Asn	Asp	Ser	Glu	Gly 150	Thr	Met	Glu	Arg	Gly 155	Glu	lle	Lys	Asn	Cys 160
Ser	Phe	Asn	Ile	165	Thr	Ser	Ile	Arg	Asp 170	Glu	Val	Gln	Lys	Glu 175	Туг
Ala	Leu	Phe	1 yr 180	Lys	Leu	Asp	Val	Val 185	Pro	Ile	Asp	Asn	Asn 190	Asn	Thr
Ser	Tyr	Arg 195	Leu	lle	Ser	Cys	Asp 200	Thr	Ser	Val	Ile	Thr 205	Gln	Ala	Cys
Pro	Lys 210	Ile	Ser	Phe	Glu	Pro 215	Ile	Pro	Ile	His	7 yr 220	Cys	Ala	Pro	Ala
225	Phe	Ala	1 l e	Leu	L ys 230	Cys	Asn	Asp	Lys	Thr 235	Phe	Asn	Gly	Lys	Gly 240
Pro	Cys	Lys	Asn	Val 245	Ser	Thr	Val	Gln	Cys 250	Thr.	His	Gly	ile	Arg 255	Pro
/al	Val	Ser	7hr 260	Gln	Leu	Leu	Leu	Asn 265	Gly	Ser	Leu	Ala	Glu 270	Glu	Glu
		275		Ser			280					285			
	290			Glu		295					300				
305					310					315					320
			•	Glu 325					330					335	
			340	Asp				345					350		
		355					360					365			Trp
	370					375					380				Thr
85					390					395					Gly 400
				405					410					415	
lsn	He		Gly 420	Leu	Leu	Leu	Thr	Arg 425	Asp	Gly	Gly	He	Asn 430	Glu	Asn

Gly	Thr	Glu 435		. Phe	e Arg	Pro	Gly 440		Gly	Asp	Het	Arg 445	Asp	Asn	Тгр	
Arg	Ser 450		Leu	ı Tyr	Lys	1yr 455		Val	Val	Lys	11e 460	Glu	Pro	Leu	Gly	
vai 465		Pro	Thr	Lys	λla 470	Lys	Arg	Arg	.Va t	Val 475	Gln	Arg	Slu	Lys		
(2)	INF	ORMA	TION	FOR	SEQ	ID	NO:2	5:								
	•	(A) L B) T C) S D) T	ENGT YPE: TRAN OPOL	H: 1 TUC DEDN OGY:	CTER 571 leic ESS: lin	base acid sing ear	pai d gle								
	(ii) MO	LECU	LE T	YPE:	DNA	(gei	nomi	c)							
	(ix	(B) L	AME/ OCAT		CDS 1		:								
	(xi) SE	QUEN	CE D	ESCR	IPTIC) : K	SEQ :	ID NO	25:	:					
ATG Met 1	GAT Asp	GCA Ala	ATG Met	AAG Lys 5	AGA Arg	GGG Gly	CTC Leu	TGC Cys	TGT Cys 10	GTG Val	CTG Leu	CTG Leu	CTG L e u	TGT Cys 15	GGA Gly	48
				Ser		AGC Ser										9 6
						TTG Leu										144
						ACC Thr 55										192
GCA Ala 65	TAT Tyr	GAT ASP	ACA Thr	GAG Glu	GTA Val 70	CAT His	AAT Asn	GTT Vel	TGG Trp	GCC Ala 75	ACA Thr	CAT His	GCC	TGT Cys	GTA Val 80	240
						CAA Gln									GAA Glu	288
AAT Asn	TTT Phe	AAC Asn	ATG Met 100	TGG Trp	AAA Lys	AAT Asn	Asp	ATG Met 105	GTA Val	GAA Glu	CAG Gln	ATG Met	CAT His 110	Glu	GAT ASP	336
ATA Ile	ATC Ile	AGT Ser 115	TTA Leu	TGG Trp	GAT ASP	CAA Gln	AGC Ser 120	CTA Leu	AAG Lys	CCA Pro	TGT Cys	GTA Val 125	AAA Lys	TTA	ACC Thr	384
CCA Pro	CTC Leu 130	TGT Cys	GTT Val	AGT Ser	TTA Leu	AAG Lys 135	TGC Cys	ACT Thr	GAT Asp	TTG Leu	GGG Gly 140	AAT Asn	GCT	ACT Thr	AAT Asn	432
ACC Thr 145	AAT Asn	AGT Ser	AGT Ser	AAT Asn	ACC Thr 150	AAT Asn	AGT Ser	AGT Ser	AGC Ser	GGG Gly 155	GAA Glu	ATG Met	ATG Met	ATG	GAG Glu 160	480
AAA Lys	GGA Gly	GAG Glu	ATA Ile	AAA Lys	AAC Asn	TGC Cys	TCT Ser	TTC Phe	AAT Asn	ATC Ile	AGC Ser	ACA Thr	AGC Ser	ATA	AGA BTA	528

80

			165	;			170			175		
		G CAG G Lr 180	Lys				Phe					576
		eet Asn				Tyr						624
	He	ACA Thr										672
His		TGT Cys			Ala							720
		AAT Asn										768
		GGA Gly 260										816
		GCA Ala										864
		AAA Lys										912
		AGA Arg										960
		GGG Gly										1008
		CAT His 340										1056
		GCT Ala										1104
		AAG Lys										1152
		TGT Cys	Gly									1200
											Asn	1248
	Glu									Lys	CAA Gln	12 96
					Glu						Pro	1344

7			*		3
					,
	•				
		• ,			

AGC Ser 450										1392
AGA Arg										1440
GGA Gly										1488
GTA Val										1536
 AGA Arg				TGA	GCG	G CC	GC			1571

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 522 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15 Ala Val Phe Val Ser Pro Ser Gln Glu Ile His Ala Arg Phe Arg Arg 20 25 30 Gly Ala Arg Thr Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro 35 40 45 Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 55 60 Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 . 70 75 80 Pro Thr Asp Pro Asn Pro Gln Glu Val Val Leu Val Asn Val Thr Glu 85 90 95 Asn Phe Asn Met Trp Lys Asn Asp Met Val Glu Gln Met His Glu Asp 100 105 110 Ite Ite Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr

Pro Leu Cys Val Ser Leu Lys Cys Thr Asp Leu Gly Asn Ala Thr Asn 130 135 140

Thr Asn Ser Ser Asn Thr Asn Ser Ser Ser Gly Glu Met Met Glu 145 150 155 160

Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg 165 170 175

Gly Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Ile 180 185 190

Pro Ile Asp Asn Asp Thr Thr Ser Tyr Thr Leu Thr Ser Cys Asn Thr

		19	5				200)				205			
Sei	- Va 210		e Th	r Gli	n Ala	215		Lys	: Val	Ser	Phe 220	Glu	Pro	He	Рго
11 <i>6</i> 225		s Ty	r Cy	s Ala	230		Gly	/ Phe	Ala	11e 235	Leu	Lys	Cys	Asn	Asn 240
Lys	Thr	Pho	e Ası	7 Gly 249	/ Thr	Gly	Pro	Cys	1hr 250		Val	Ser	Thr	Val 255	Gln
Cys	Thr	His	260 260		Arg	Pro	Val	Val 265		Thr	Gln	Leu	Leu 270	Leu	Asn
Gly	Ser	275		Gli	Glu	Glu	Val 280		lle	Arg	Ser	Ala 285	Asn	Phe	Thr
Asp	Asn 290		Lys	Thr	Ile	I l e 295	Val	Gln	Leu	Asn	Gln 300	Ser	Val	Glu	Ile
Asn 305	Cys	Thr	Arg	Pro	310	Asn	Asn	Thr	Arg	Lys 315	Ser	He	Arg	Ile	Gln 320
Arg	Gly	Pro	Gly	Arg 325	Ala	Phe	Val	Thr	11e 330	Gly	Lys	Ile	Gly	Asn 335	Met
Arg	Gln	Ala	His 340		Asn	lle	Ser	Arg 345	Ala	Lys	Trp	Asn	Ala 350	Thr	Leu
Lys	Gln	1 l e 355	Ala	Ser	Lys	Leu	Arg 360	Glu	Gln	Phe	Gly	Asn 365	Asn	Lys	Thr
lle	1 l e 370	Phe	Lys	Gln	Ser	Ser 375	Gly	Gly	Asp	Pro	Gl u 380	He	Val	Thr	His
Ser 385	Phe	Asn	Cys	Gly	Gly 390	Glu	Phe	Phe	Туг	Cys 395	Asn	Ser	Thr	Gln	Leu 400
Phe	Asn	Ser	Thr	1 rp 405	Phe	Asn	Ser	Thr	1rp 410	Ser	Thr	Glu	Gly	Ser 415	Asn
Asn	Thr	Glu	Gly 420	Ser	Asp	Thr	He	1hr 425	Leu	Pro	Cys	Arg	1 l e 430	Lys	Gln
Phe	He	Asn 435	Met	Val	Gln	Glu	Va l 440	Gly	Lys	Ala	Met	Tyr 445	AlB	Pro	Pro
lle	Ser 450	Gly	Gln	Ile	Arg	Cys 455	Ser	Ser	Asn	Ite	1hr 460	Gly	Leu	Leu	Leu
1hr 465	Arg	Asp	Gly	Gly	Asn 470	Asn	Asn	Asn	Gly	Ser 475	Glu	Ile	Phe	Arg	Pro 480
Gly	Gly	Gly	Asp	Me t 485	Arg	Asp	Asn	Trp	Arg 490	Ser	Glu	Leu	Tyr	Lys 495	Туг
Lys	Val	Val	Lys 500	Ile	Glu	Pro	Leu	Gly 505	Val	Ala	Pro	Thr	Lys 510	Ala	Lys
Arg	Arg	Val 515	Val	Gln	Arg	Glu	Lys 520								

- (2) INFORMATION FOR SEQ ID NO:27:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1532 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..1522
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

	(xi) SE	QUEN	ICE C	ESCR	IPTI	ON:	SEQ	ID N	0:27	:					
	Asp								TGT Cys 10	Val						48
GCA Ala	GTC Val	TTC	GTT Val 20	Ser	Pro	AGC Ser	CAG Gln	GAA Glu 25	ATC	CAT His	GCC Ala	CGA Arg	TTC Phe 30	AGA Arg	AGA Arg	96
			Val						ACA Thr							144
		Lys							TŢT Phe							192
									TGG Trp							240
									GTA Vai 90							288
									GTA Val							336
									AAG Lys							384
									GAT Asp							432
									AGA Arg							480
									GAT ASP 170							528
									CCA Pro					Asn		576
									TCA Ser						TGT Cys	624
															GCT	672
															GGA Gly 240	720

			AAT Asn		Ser					Thr						768
-			ACT Thr 260	Gir												816
			AGA Arg													864
GTA Val	CAG Gln 290	Leu	AAA Lys	GAA Glu	TCT Ser	GTA Val 295	GAA Glu	ATT Ile	AAT Asn	TGT Cys	ACA Thr 300	AGA Arg	CCC Pro	AAC Asn	AAC Asn	912
AAT Asn 305	Thr	AGA Arg	AAA Lys	AGT Ser	ATA Ile 310	His	ATA Ile	GGA Gly	CCA Pro	GGG Gly 315	AGA Arg	GCA Ala	TTT Phe	TAT Tyr	ACT Thr 320	960
			ATA Ile													1008
			TGG Trp 340													1056
			GAG Glu													1104
			ATT lle													1152
			TCA Ser													1200
			AAT Asn													1248
			ATT Ile 420				Val									1296
						Gln									GGG	1344
			ACA Thr		Asp										ATC	1392
TTC Phe 465	AGA Arg	CCT Pro	GGA Gly	Gly	GGA Gly 470	GAT Asp	ATG Met	AGG Arg	GAC Asp	AAT Asn 475	TGG Trp	AGA Arg	AGT Ser	GAA Glu	TTA Leu 480	1440
TAT Tyr	aaa Lys	TAT Tyr	Lys	GTA Val 485	GTA VBl	AAA Lys	ATT Ile	Glu	CCA Pro 490	TTA Leu	GGA Gly	GTA Val	GCA Ala	CCC Pro 495	ACC	1488
AAG Lys	GCA Ala	Lys	AGA Arg 500	AGA Arg	GTG Val	GTG Val	Gln	AGA Arg 505	GAA Glu	AAA Lys	TG	AGCG	GCCG	C	•	1532

(2) INFORMATION FOR SEQ ID NO:28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 507 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: Linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ 1D NO:28:

Met Asp Ala Met Lys Arg Gly Leu Cys Cys Val Leu Leu Leu Cys Gly 1 5 10 15

Ata Val Phe Val Ser Pro Ser Gln Glu Ile His Ata Arg Phe Arg arg 20 25 30

Gly Gly Arg Val Glu Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro
35 40 45

Val Trp Lys Glu Ala Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys 50 60

Ala Tyr Asp Thr Glu Val His Asn Val Trp Ala Thr His Ala Cys Val 65 70 75 80

Pro Thr Asp Pro Asn Pro Gin Glu Val Val Leu Glu Asn Val Thr Glu 85 90 95

His Phe Asn Met Trp Lys Asn Asn Met Val Glu Gln Met Gln Glu Asp 100 105 110

lle Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr 115 120 125

Pro Leu Cys Val Thr Leu Asn Cys Lys Asp Val Asn Ala Thr Asn Thr 130 135 140

Thr Asn Asp Ser Glu Gly Thr Met Glu Arg Gly Glu Ile Lys Asn Cys 145 150 155 160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asp Glu Val Gln Lys Glu Tyr 165 170 175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asn Asn Thr 180 185 190

Ser Tyr Arg Leu Ile Ser Cys Asp Thr Ser Val Ile Thr Gln Ala Cys 195 200 · 205

Pro Lys Ile Ser Phe Glu Pro Ile Pro Ile His Tyr Cys Ala Pro Ala 210 215 220

Gly Phe Ala Ile Leu Lys Cys Asn Asp Lys Thr Phe Asn Gly Lys Gly 225 230 240

Pro Cys Lys Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Arg Pro 245 250 255

Val Val Ser Thr Gin Leu Leu Leu Asn Gly Ser Leu Ala Glu Glu Glu Glu 260 265 270

Val Val Ile Arg Ser Asp Asn Phe Thr Asn Asn Ala Lys Thr Ile Ile 275 280 285

Val Gln Leu Lys Glu Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn 290 295 300

Asn Thr Arg Lys Ser Ile His Ile Gly Pro Gly Arg Ala Phe Tyr Thr

The Siv Giu Ite Ite Gly Asp Ite Arg Sin Ata His Cys Ash Ite Ser 325

Arg Ata Lys Trp Ash Asp Inn Leu Lys Gin Ite Val Ite Lys Leu Arg 340

Siu Gin Phe Glu Ash Lys Inn Ite Val Phe Ash His Ser Ser Gly Gly Asp Pro Glu Ite Val Met His Ser Phe Ash Cys Gly Gly Glu Phe Phe 370

Tyr Cys Ash Ser Thr Gin Leu Phe Ash Ser Thr Trp Ash Ash Ash Inn Ang Asp Glu Gly Ser Ash Ash Thr Glu Gly Ash Thr Ite Thr Leu Pro Cys Arg 405

Ite Lys Gin Ite Ite Ash Met Val Gly Ash Thr Ite Thr Leu Pro Cys Arg 405

Ata Pro Pro Ite Arg Giy Gin Ite Arg Cys Ser Ser Ash Ite Thr Gly 435

Leu Leu Leu Thr Arg Asp Gly Gly Ite Ash Glu Ash Gly Thr Glu Ite 455

Tyr Lys Tyr Lys Val Val Lys Ite Glu Pro Leu Gly Val Ata Pro Thr Lys Arg Arg Arg Arg Val Val Glu Pro Leu Gly Val Ata Pro Thr Lys Arg Arg Arg Val Val Glu Pro Leu Gly Val Ata Pro Thr Lys Arg Arg Arg Val Val Glu Arg Glu Lys

Sob

- (2) INFORMATION FOR SEQ ID NO:29:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 15 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
 - Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg 1 5 15

What is claimed is:

- A recombinant nucleic acid molecule which encodes a mutant HIV-1 gp120 envelope glycoprotein comprising a
 V3 loop deletion and a C4 domain_(W->X) point mutation, wherein X is an amino acid residue other than tryptophan.
- The recombinant nucleic acid molecule of claim 1,
 wherein X is a valine residue.
 - 3. The recombinant nucleic acid molecule of claim 1, wherein the nucleic acid molecule is a DNA molecule.
- 15 4. The recombinant nucleic acid molecule of claim 3, wherein the DNA molecule is a plasmid.
- 5. The recombinant nucleic acid molecule of claim 4, wherein the plasmid comprises the sequence of the plasmid designated PPI4-tPA.
 - 6. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1_{IAI} gp120 envelope glycoprotein C4 domain.

25

- 7. The recombinant nucleic acid molecule of claim 6, wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1_{IAI} gp120 envelope glycoprotein.
- 30 8. The recombinant nucleic acid molecule of claim 1, wherein the C4 domain is an HIV-1 $_{\rm IR-FL}$ gp120 envelope glycoprotein C4 domain.
 - 9. The recombinant nucleic acid molecule of claim 8,

WO 94/22477

15

25

30

wherein the mutant HIV-1 gp120 envelope glycoprotein is a mutant HIV-1 $_{\rm JR-FL}$ gp120 envelope glycoprotein.

- 10. The mutant HIV-1 gp120 envelope glycoprotein encoded by the recombinant nucleic acid molecule of claim 1.
 - 11. A vaccine which comprises a therapeutically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.

12. A method of treating an HIV-1-infected subject, which comprises immunizing the HIV-1-infected subject with the vaccine of claim 11, thereby treating the HIV-1-infected subject.

13. A vaccine which comprises a prophylactically effective amount of the mutant HIV-1 gp120 envelope glycoprotein of claim 10, and an adjuvant.

20 14. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the HIV-1-exposed subject's becoming infected with HIV-1.

15. A method of reducing the likelihood of a non-HIV-1-exposed subject's becoming infected with HIV-1, which comprises immunizing the non-HIV-1-exposed subject with the vaccine of claim 13, thereby reducing the likelihood of the non-HIV-1-exposed subject's becoming infected with HIV-1.

16. A method of obtaining partially purified antibodies which specifically bind to the CD4-binding domain of •

WO 94/22477

5

30

PCT/US94/03282

89

HIV-1 gp120 envelope glycoprotein, which method comprises (a) immunizing a non-HIV-1-exposed subject with the vaccine of claim 13, (b) recovering from the immunized subject serum comprising said antibodies, and (c) partially purifying said antibodies, thereby obtaining partially purified antibodies which specifically bind to the CD4-binding domain of HIV-1 gp120 envelope glycoprotein.

- 10 17. The method of claim 16, wherein the subject is a human.
 - 18. The partially purified antibodies produced by the method of claim 16.
- 15 19. A pharmaceutical composition, which comprises a therapeutically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.
- 20 20. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1-infected cells in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
 - 21. A method of treating an HIV-1-infected subject, which comprises administering to the subject a dose of the pharmaceutical composition of claim 19 effective to reduce the population of HIV-1 in the HIV-1-infected subject, thereby treating the HIV-1-infected subject.
- 22. A composition which comprises a prophylactically effective amount of the partially purified antibodies of claim 18, and a pharmaceutically acceptable carrier.

5

- 23. A method of reducing the likelihood of an HIV-1-exposed subject's becoming infected with HIV-1, which comprises administering to the HIV-1-exposed subject a dose of the composition of claim 22 effective to reduce the population of HIV-1 in the HIV-1-exposed subject, thereby reducing the likelihood of the subject's becoming infected with HIV-1.
- 10 24. The method of claim 23, wherein the subject is a medical practitioner.
 - 25. The method of claim 23, wherein the subject is a newborn infant.

15

- 26. A method of reducing the likelihood of a non-HIV-1exposed subject's becoming infected with HIV-1 as a
 result of exposure thereto during an incident wherein
 there is an increased risk of exposure to HIV-1, which
 comprises administering to the subject immediately
 prior to the incident a dose of the composition of
 claim 22 effective to reduce the population of HIV-1 to
 which the subject is exposed during the incident,
 thereby reducing the likelihood of the subject's
 becoming infected with HIV-1.
 - 27. The method of claim 26, wherein the subject is a medical practitioner.

FIGURE 1

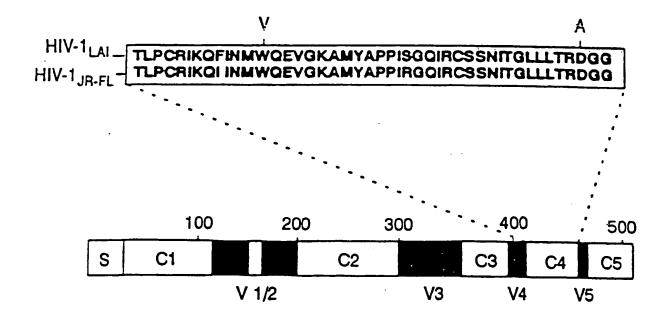
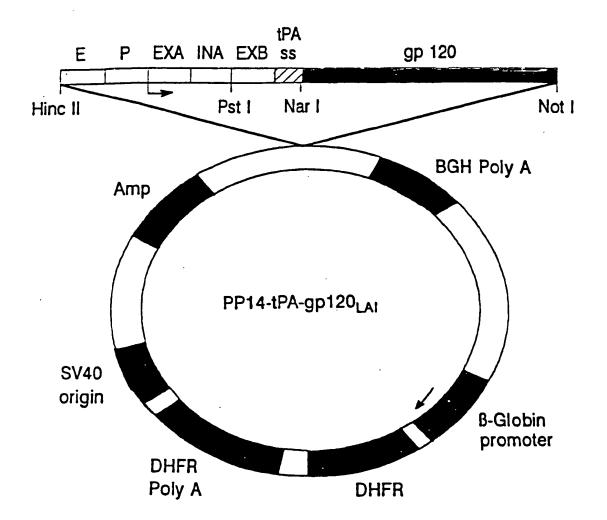


FIGURE 2



acggggatttccaagtctccacccattgacgtcaatgggaytttgttttg@caccaaaatcaacgggactt

433

3/42

FIGURE 3A FIGURE 3B FIGURE 3D FIGURE 3E FIGURE 3F	ittacggggtcattagttcatagcccatatatgga	ict tacggtaaatggeeegeetggetgaeegeeeaaegaeeeeegeeeattgaegte	ctttccattgacgtcaatgggtggactatttacg	tgccaagtacgcccctattgacgtcaatgacgg	tatgggacttcctacttggcagtacatctacgt	accatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactc
FIGURE 3A	HincII ttgac attgattattgactagttattaatagtaatcaattacggggtcattagttcatagcccatatatgga	gttccgcgttacataacttacggtaaatggcccgcctg	aataatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggactatttacg	gtanactgeceaettggeagtacateaagtgtateatatgeeaagtaegeeeetattgaegteaatgaegg	taaatggcccgcctggcattatgcccagtacatgaccttatgggacttcctacttggcagtacatctacgt	attagtcatcgctattaccatggtgatgcggttttggc
	-	73	145	217	289	361

SUBSTITUTE SHEET (RULE 26)

FIGURE 3B

tccaaaatgtcgtaacaactccgccccattgacgcaatgggcggtaggcytgccgttgcggoggtgggaggtetatat

505

acttaggcacaggacaatgcccaccaccagtgtgtgcgcacaagggccgtggcggtagggtagggtatgtctga a agcagaget egt t tagt gaa ee GTCAGATCGCCTGGAGACGCCATCCACKTGTTTTGACCTCCATAGAAG ACACCGGGACCGATCCAGCCTCCGCGGCCGGGAACGGTGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGA attggetatatgecaatactetgteetteagagaetgaeaeggaetetgtatttaeaggatgggtees Cgtaagtaccgcctatagactctataggcaccctttggctctttaggctcttatgctatgctatactgttttggcttg ggccaacacccgtcctagataggtgatggtatacttagcctataggtgtgtgggtgttggttattgaccattat cactecetattggtgaegataettteeattaetaateeatgaegeegetettgeeaeaetatet Intron A Exon A 577 649 1009 865 1153 793 721 937 1081 1225

SUBSTITUTE SHEET (RULE 26)

FIGURE 3C

gaccccarccarababgtagtattegtaatgtcacabaaattttancatgtcbaaaatgtgt tgtgcatcagatgctaaagcatatgatacagaggtacataatgtttgggccacacatgcctgtgtaccaca abatgageteggagattgggetegeacgetgaegeagatggaagaettaaggeageageggaagaaga GTGCTGCTGTTGTGGAGCAGTCTTCGCCCAGCCAGGAAATCCATGCCCGATTCAGAAGAGGGCGCC aggcagctgagttgtattetgtagagttggaggtaactcccgttgcigttgctgttaacggtgct teettteeatgggtettteeggagicaccitetitacaccaitgaafeaaigaaggetettee gtgtagtetgageagtaetegttgctgccgcgcgcgcaccagaetaataatagctgacagaeagactgacagactaacagactg NarI tPA signal sequence œ æ ⋖ € -2 23 × 33 ga, ⋗ w Z Q ۵ ပ മ Exon B [c] **ම** Z ≫ H u. 0 **ہ**ے PstI ≫ Ø ⊱ ۶ ⋖ Œ △ Signal cleavage ര 25 (A) Œ ധ 0 শ্ৰ ഗ শ্ৰ 1729 1657 1801 1585 1369 1513 1297 1441

5/42

SUBSTITUTE SHEET (RULE 26)

FIGURE 3D

CTAGCAG/IAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTG SEGGAAATGATGATGGAGAAAGGAGAGATAAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG GTGCAGAAAGAATATGCATTTTTTTATAAACTTGATATAACAATAGATAATAGATAATACTACCAGCTATACG TATTGTGCCCCGGCTGGTTTTGCGATTCTAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACA aatgtcagcacagtacaatgtacacatggaattaggccagtagtatcaac'tcaactgctgttgaatggcagt TIGACAAGTIGTAACACCTCAGTCATTACACAGGCCTGTCCAAAGGTATCCTTTGAGCCAATTCCCATACAT TGTGTTAGTTTAAAGTGCACTGATTTGGGGAATGCTACTAATACCAATAGTAGTAATACCAATAGTAGTAGT GAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCATGTGTAAAATTAACCCCACTC S U S ø H S 4 .1 Z G H × ഗ H 1 Δ ပ H Z 回 S) Z O z ပ ဟ Ĺ., S Δ, S × z Z H > S Z H × H بعا م × > 0 S Z တ Z 0 H U H ပ Z Ω 4 ပ z Ω 4 œ z z 3 × × .. 0 ၒ -1 ... × H ပ H ഗ .1 H I L **>** S Œ Ω ~ H G H ပ H × ഗ Ĺ > <u>[a</u> Δ ပ ပ 0 M > X X × > (L) K ы .1 W ပ H I L) 4 က I တ ဟ u > L > ပ ပ 2305 2089 2161 2377 1873 1945 2233 107 2017 155 179 203 227 251 275 131

6/42

SUBSTITUTE SHEET (RIII F 26)

GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA gaaggaagtgacacaacactcccatgcagaataaacaatttataaacatgggcaggaagtaggaaa gcaatgtatgcccctccatcagcggacaaattagatgttcatcaaatattacagggctgctattaacaaga tgtrattcaacrcaactgtttratrctrcttgctttratrctacttccactactccact · CCAGGGAGAGCATTTGTTACAATAGGAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA aantggnatgccactttaaaacacatagctagcaaattaagagaacaatttggaaataataaacaataatc TTTAAGCAATCCTCAGGAGGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTTAAC aaccaatctgtagalltaattgtacaagacccaacaatacaagaagaaaaagtatccgtatccagaggga × ပ O W ഗ ල ල 0 Ø, ပ 33 X. M G ပ E 0 ဟ Ę ഗ ග Z Z [24 幺 0 0 33 ပ \bowtie Z œ ပ ø ഗ H (<u>17</u> ഗ <u>[1]</u> ڪ 0 က ø æ ഗ Œ Z ပ œ ٦, (F Z 25 Z Z Œ D ල 쏪 యి **~** ഗ 33 O Æ U 24 **[13**] **(4)** ø œ ပ ပ ഗ **~** ھ مھ **~** ഗ ග 0 Q Z U ⊶ ၒ Z 幺 Z ဗ Q۵ **,**2 Z **₽** Þ ഗ 0 Q, Z **Ç**~3 (M) (C2) æ Z S Q ß Œ × ഗ 0 ල X 23 2809 2953 439 883 467 2889 2593 2665 2737 395 2881 323 2521 347 371

7/42

CHECTITHE CHEET /RIH F 26)

FIGURE 3F

æ G GTGGTGCAGAGAAAAATGAGCGCCGC NotI æ . ы 3025 3097

SUBSTITUTE SHEET (RULE 26)

FIGURE 4

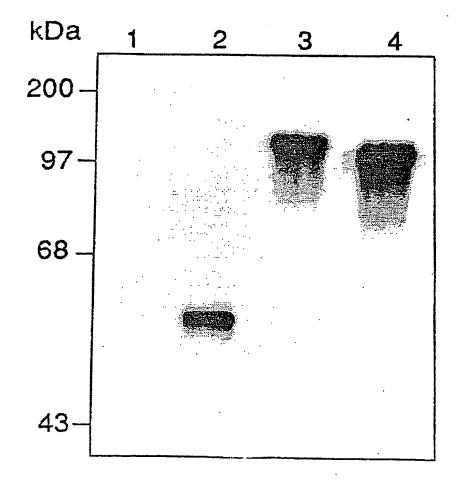


FIGURE 5A

[gp120]
(ng/ml)
6
14
123
4
18
18

FIGURE 5B

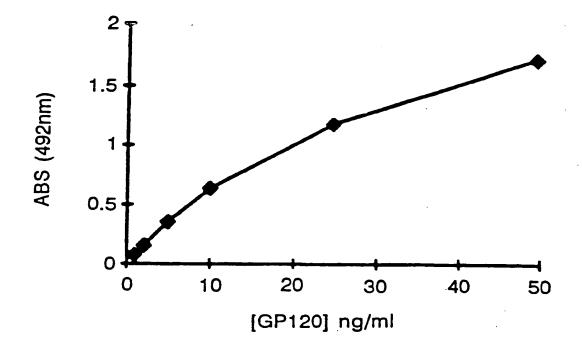


FIGURE 6

kDa 1 2 3 4 200-97-68-

29-

43-

ttaaccccactgtgttactttaattgcaagatgtgaatgctactactaataccactaat

319

107

379

127

23

⋖

Z

Q

幺

Z

ವಿ

gatacccacccalccaticacacacacacaltaaaaaactccttttcaatataccaca

Z

ശ

ပ

Z

25

⋈

പ്ര

ල

Œ

ල

(M)

839

147

JR-FL

9

79

27

139

97

13/42

Caagaagtagtattggaaaatgtaacagaacattttaacatggaaaaataacatggta CATGCCCGATTCAGAAGA@@@@@AGAGTAGAAAAGTTGTGGGTCACAGTCTATTATGGG GTACCTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT GATACAGAGGTACATAATGTTTGGGCCACATGCCTGTGTACCCACAGACCCCAAACCC atggaticcaatgaagaga gaacagatecaggatataatatattataataaaagcctraagccaaagtaaaa ල C Z FIGURE 7C FIGURE 7A FIGURE 7B ⋖ ⋖ ഗ Z Q Q ٥ ഗ ഞ 83 **@** ⋖ △ Signal cleavage 23 Ð ശ ပ C L 23 O FIGURE 7A وكئ Œ ß **Q** ھے ⋖ æ Œ 23 ල **(** <u>[73</u>] ~3 ധ ഭ ß ⋖ ⊱ Ŋ NarI Ø œ Ð Ξ Ŋ ල Æ ð Z ⋈ প্র Œ æ 网 ശ്ര Q Ø 出 Œ ھ ശ്ര ധ ۱L 23 A 0 ð ပ Œ Œ <u>M</u> D 0 ⋖ শ্ৰ (ca)

SUBSTITUTE SHEET (RULE 26)

199

67

259

8

@ <u>/</u>
E F
الالح

899	accataacactgcacaaacaatatctcttttttataacttgatgtacca
167	s i r d e v q k e y a l f y k l d v v p
55.9 187	ATACATAATAATACCAGCTATAGTTGATAAGTTGTGACACCTCAGTCATTACACAG
619	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCGGCTGGTTTT
207	A C P K I S F E P I P I H Y C A P A G F
679	GCGATTCTAAAGTGTAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC
722	A I L K C N D K T F N G K G P C K N V S
739 785	acagtacaatgracaattaggccagtagtatcaactcaac
799	agtctagcagaagaggagtagtaattagatctgacaatttcacgaacaatgctaaaacc
267	s la e e e e v v i r s d n f t n n a k t
8 8 7	atantagtacagctgaarctgtagaaattaattgtacaagacccaacaacaataca i i v q l k e s v e i n c t r p n n n t
918	agraragtatacatataggaccagggaggattttatacagggagaataatagga
307	r k s 1 h 1 c p g r a f y t t g e 1 i g
979	Gatrtragacacatttgtaacattagtagacaaaatggaatgrtataaaacag
327	d i r q a h c n i s r a k w n d t l k q

FIGURE 7C

ATAGTTATAAAATTAAGACAATTTGAGAATAAACAATAGTCTTTAATCACTCCTCA

0

1039

ഗ

660	& 28€	88	SACC	SCAC	BARA	TTG	TM	360	ACA Cal	FFT	TAA	rrei	B	GGAGGGGACCCAGAAATTGTAATGCACAGTTTTAATTGTGGAGGAGAATTTTTCTACTGT	GAB	TLL	rrci	S&C	TCL	
367	ဖ	ဖ	Q	<u>@</u>	(Sa)	~ 3	≫	>E	96 96		Z	U	O	ဖ	[62]	[1 ₂	دما	≫	ပ	
159	A.R.T.	$\mathbb{I}^{\mathcal{C}M}$	SCA SCA	CARAC	TCI	TTB	ATA	GTA(FFC	SGA.	TAR	raan	C.	aa itcaacacaactgittaatagtacttggaataataatactgaagggtcaaataacact		₹Ç.	AAT!	C	*CT	
387	Z	S	⊱	ō	~ 3	ගිය	æ	S	8 <u>~</u> 3	%	23	Z	S)	O LF N S T M N N T E G S N N	ග	ശ	Z	z	e-	
219	G & A	GRAGGRAA	aa T	acta	TC3	CAC	J L	CAT(3C&C	M	RARA	ACAL	W.F.	.tactatcacactccatgcagaataaaacaaataaaacatgtgccaggaa	A C	at C	ည္သင္သ	છ	8	
407	M	ဖ	22	(⊷	⊢	⊱	بَ	<u>o</u> .	es.	⊶ ~	25	O	×	⊶	Z	E	33	Ō	ហ	
279	GTB	BG	RABA	3C &	\TGT	at	CCC	CIC(CAI	الكي	ßGG!	ACAL	WY.	gtaggaaagcaatgtatgccctcccatcagaggacaaattagatgtcatcaaatatt	'ÎGT	TCA	TC&	AAT	ATT	
827	۵	ဗ	×	8		≫	æ	<u>~</u>	D.	6 5.	ල ්	Ö	3	Y A P P I A G Q I A C S S N I	ပ	ഗ	ശ	Z	⋈	
3 30	AC A	3000	CIG	crai	TA		S S S S	a Ta	376	Ty	TAR	rga(SAA.	a caccectect and that carea cate of the tratea carea color careaters.	SCC.	G ^{ନ୍ଦ}	arc.	rrc	aga A	
887	E-	ဗ	ا ا	2	Z)	y ₩	Œ	<u></u>	w w		23	ശ്ര	Z	LLTRDGGINENGTEIF	æ	ശ്രീ	⊶	(Es)	Ø\$	
000					R. T. B.	TGA	වුදුව	BCA.	att.	3GAG	Sy C	rca	\FT!	o pre cer cer cer cer cer cer compare	S. S	THL	MAN.	STA	GTA	
&67 &67) ()	ခြို့ဖ	ြို့မ	ල	Q	Œ	OS.		92. 93.	60. 82	S)	ശ്ര	2	C C D M R D N M R S E L Y R Y K V	25	₩	25	∌	A	
85.9	A.A.A.	att(GAM	3 C AI	Frag	SEAG CARC	Tag	CAC	3	S. C.	7399)) 	3AG	aaaattgaaccattagcactccaccaagcaaagagaaaagagagaa	G T6	GTG		3G.	G M	
487	25	×	<u>(2)</u>	<u>@</u>	শ্ৰ	ල	>	 &	6.	25 ~	Œ	25	氐	B T C V A P T K A K A W W V O M	≫	≫	0	<u>o</u> \$	<u>M</u>	
			2	NotI															•	
519	A.A.A.	TGM		aaa toarcorooco	Ç															
507	×																			

FIGURE 8C FIGURE 8A FIGURE 8B

FIGURE 8A

LAI AV3

ATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTG U æ ഠ øŝ ¥ æ ⋖

CTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGGAAATCCATGCCCGATTCAGAAGAGGGGGCGAGACA Nar ණ ණ œŝ ⋖ I ⋈ w G ඟ ത ∌ ⋖ **©** U æ 37

GALARATIGIGGERCACAGICIRITATGGGIACCTGTGTGGAAGGAAGCAACCACCACCACTCTATTTGTGCA Signal cleavage 🛆 ര

TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCC æ W 26 83 ∌ **©** ≫ ≫ \approx Ð 23 ھ 109 181 37

aacccacaagaagtagtattggtaaatgtgacagaaaattttaacatgtggaaaaatgacatgtaacag Œ ပ Z Æ 25 X \mathfrak{Z} H Œ Œ Z 33 وحي Z Z (J 靐 Ç ð ⊳ (ca) Z Þ H ~3 Ω ≫ D Ð Œ 26 W æ Ω တ 253 61

0

85

at gcat cagcatat mat cagtitat gggat caagcctaaagccatgt aanattanceccactettett ç ھے 25 Ð ပ <u>@</u> 25 **~**3 ഗ 0 Q 23 ~7 ഗ 325 109 acttraagtecactgatttggggaatgctactaataccaatagtagtaataccaatagtagtagtgggggaa W 397

G S ഗ ഗ Z ۲ Z ശ (J) Z Z ⋖ Z G 133

AGCACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGTTGAATGGCAGTCTAGCA **GAAGAAGAGGTAGTAATTAGATCTGCCAATTTCACAGACAATGCTAAAACCATAATAGTACAGCTGAACCAA** ATGATGATGGAGAAAGGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAGGTGCAG GCCCCGGCTGGTTTTGCGATTCTAAAATGTAATAATAAGACGTTCAATGGAACAGGACCATGTACAAATGTC tctgtagalattaattgtacaggtgctggacattgtaacattagtagggcalaatggaatgccactttaaa **aaagaatatgcatttttttataacttgatataataccaatagatatgatatc**cagctatacgttgaca <u>agttgtaacacctcagtcattacacaggcctgtccaaaggtatcctttgagccaattcccatacattattgt</u> ¥ × O G **>** S ۵, æ Z H G H H လ H ۵, H H H Ω H [L] <u>ෆ</u> H O × z z **~** ഗ [a, Δ K H တ က H بعا H တ Z z > H Ω Ĺ × Z ۵, ¥ > H > U z လ H ۵, ۵, <u>[_,</u> ပ H z 四 0 O æ z Z ,, 4 4 × 0 × ပ S U -1 H æ **= >** M U ~ H Н G Ĺ ပ > Z × بعا 0 G M 4 W X I (L) u 829 277 901 685 229 757 253 301 613 469 205 157 541 181

SUBSTITUTE SHEET (RIII F 26)

FIGURE 8C

CAGATAGCTAGCAAATTAAGAGAACAATTT**GGAAATAATAAAACAATAA**TCTTTAAGCAATCCTCAGGAGGG **GGGTCCGAGATCTTCAGACCTGGAGGAGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAA** CICCCATGCAGAATAAAACAATTTATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC **AGCGGACAAATTAGATGTTCATCAAATATTACAGGCTGCTATTAACAAGMGATGGTGGTAATAACAA**CAAT GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACACTGTTT Z W S æ G ഗ H O S 0 G × ပ Σ ы Z **∵** ⊶ K <u>م</u> (e, တ Ы > . [→ × Œ Н æ H H O Z Z æ H H z > z × بعا .-Ω × (L) ဟ ы ~ . Z G Œ ပ 0 × Z ပ Σ ပ 3 ы H ပ Ω ပ Σ H H ۵, U Ĺ z z H ഗ 4 0 U Н z S سا 3 > ပ E S H G Ĺ., Œ 0 S X ဟ م H H × H Z ပ æ ۵, œ × > H بعا Į. (L) H H S H 3 ø ပ (L) ر انت 2000000 NotI ၒ > ρ, S ഗ Ω တ G 1045 349 973 325 1189 1333 1405 1117 445 373 397 1261 421 469 1477

18/42

SUBSTITUTE SHEFT (RUI F 26)

								:
	AGA R	GGGCTCTGCTGTGTGTGTGTGAGCAGTCTTCGTTTCGCCAGCAGAAATC	Nafi Catgecegaticagaagaagaagaagaagaagaagaagagaagaga H A R F R R G A R V E K L W V T V Y Y G	GTACCTGTGGGAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT V P V W K E A T T T L F C A S D A K A Y	GATACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACCAGACCCCAACCCA	CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA Q E V V L E N V T E H F N M W K N N M V	GANCAGATGCAGGATATAATCAGTTTATGGGATCAAAGCCTALAGCCATGTGTAAAA E Q M Q E D I I S L W D Q S L K P C V K	TIAACCCCACTCTGTGTTACTTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT L T P L C V T L N C K D V N A T N T T N
	ATGGATGCAATGAAGAGA M D A M K R	S ⊒	TAT	80 4	AAON	ATG	GTA V	ACT
	ATG.	CA Q	TAT	*	ည	Z Z	ည်ပ	ACC
98 98	ზ ∢	A GC 8	GIC V	GCT •	GAC	AAT	P C	ZAZ
FIGURE 9A FIGURE 9B FIGURE 9C	GAT	ည္သ	A CA	GAT	Ž t	₹×	۲ ۲	ACT.
FIG	A A A A	කිරි සි	GTC V	TCA	ည က	7GG ₹	CTA	3CT
	·	GTT	1 જ	ක්රී ද්රී	GTA	ATG	ည်လ	AAT
		TIC	TTG	TGT	រត្តរ ក	\frac{1}{2}	80	GTG/
_		GTC	AAG ×	rrr F	ည္ဟ	TTT	CAT D	GAT(
6::		A SCA	S S	gna CTA	CAT	CAT	₹	AAG K
FIGURE 9A		ပ္ပ်ံ ပ	GTA	ACT T	Ž H	GAA E	TTA	ည်
פֿו		រក្សា	40 K	TOP:	ည္ပ	Ž to	S	AAT
u_		St 4	Nari Ogge A	₩ •	3. T.G.	\ \ \	ATO I	TTA
•		Si sa	≥႘ၟ ဗ	₹ 25	\GT7	Z Z	IATA	ACT
		SC IC	S K	15 3	\{\frac{1}{2}} \text{z}	Д ы	GAT) V
		rgro V	ZAG.	Šχ	S =	L	363 E	7. 1.01
		ည် ပ	ATTO	अस्ट भ	3GTJ	VGT.	50	CTC
		ရှိ ပ	(S) K	rgrc v	AGA(VGT/	ATC H	ည်
		GCTCTGCTGTGTGCTGCTGTGTGAGCAGTCTTCGTTTCGCCAGCAAA	Nafi Satgecegatteagaagagagagagaagaagateagateattatatgge H A R F R R G A R V E K L W V T V Y Y G	ည်ရှိ	[} T	AAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGT! Q E V V L E N V T E H F N M W K N N M V	ancagatgcaggagatataatcagtttatgggatcaaagcctalagccatgtgtaaa. E o m o e d i i s l w d o s l k p c v k) S
	<u>د</u> م	ဗွ် ဖ	E CA	GT)	GA.	30	S a	TT T
	-FL AV3							
	4	19 7	79	39	99	59 87	19	79

439	GATA	ଧୁ	SAG	1		ATG	GAG	AGA	4 5	₹8	ATA	*	M C	TGC	TCT	TIC	MAT	ATC	ACC	ACA
147	Ω	ဟ	SEGTMERGEIKNOSFNITT	ပ	۲	Σ	(L)	œ	<u>ဖ</u>	(a)	-	×	Z	ပ	လ	(LL)	z	H	⊬	H
499	AGC	AIA	AGA	IGA1	GAG	GIG	CAG	A	₽	TAT	GCT	CIT	TIT	TAT	A A	Cir	GAT	GTA	GTA	₹
	S	H	SIRDEVOKEYALFYKLDVVP	Ω	ы	>	0	×	ы	>	4	1	(Fr	>	×	'n	. Δ	>	>	۵,
559	ATA	GAT	X	M	M	300	3 6	TAT	8 66	116	ATA	AGT	TGT	GAC	ACC	TC'A	GTC	ATT	ACA	CAG
187	H	Ω	IDNNTSYRLISCDTSVITO	Z	z	H	လ	>	æ	H	H	S	ပ	Ω	H	ະທຶ	>	H	H	0
-	ပ္ပ	īĞī	GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT	3	ATA	ATCC	TTT	GAG	ర్ర	ATT	ည	ATA	SA	TAT	TGT	ည	ຽວ	GCT	GGT	TTT
207	~	ပ	ρ,	×	H	S	<u>L</u>	ш	۵,	H	۵.	н	x	>	ပ		م	4	G	Ĺ
619	ပ္ပ	ATT	CTA	3	TGI	MI	GAT	AAG	ACG	TTC	AAT	SGA G	¥	GGA GGA	CCA	TGT	AA.	AAT	GIC	AGC
~	<	H	A I L K C N D K T F N G K G P C K N V S	×	ပ	Z	Ω	×	H	(a,	Z	ပ	×	O	O.	رن ⁻	×	z	>	S
739	S	E	ACAGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTGCTAAATGGC	VIG.	Ş	3	35	ATT	3 56	Ş	GTA	GTA	TCA	ACT	Š	CTG	CTG	CIA	AAT	ည္ဟ
247	H	>	a	ပ	H	×	ပ	н	œ	۵.	>	>	လ	E	0	H	-	H	Z	ပ
700	AGT	Ç	AGTCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC	S	3	IGA G	GTA	GTA	ATT	AGA	TCT	GAC	MAT	TTC	ACG	AG	AAT	SCI	*	ACC
267	S	-	4	ш	ш	ы	>	>	H	œ	S	Ω	z	Ĺ	H	Z	Z	«	×	⊬
9	ATA	ATA	CT.	CAG	CIC	X	SA	ICI	GTA	₹	ATT	AAT	Ę	2	GGT	SCI	3	5	TGT	AAC
7 @	Н	H	IIVOLKESVEINCTGAGACN	Œ	-	×	ப	တ	>	ш	н	Z	ບ	E +	G	4	Ö	岡	ပ	Z
•	ŧ		THE CONCLOS DE DECIDE DE LEGGE DE LE CACTETARA CAGATAGETA TARA LA TERA GAGA CA CA	֓֞֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓	4 4 4 4	T.C.	TAA	ָרָאָרָי האָרָי	ACT	TTA	***************************************	CAG	ATA	GTT	ATA	ALA	TTA	AGA	8	S S
919	AII	2		<u>ک</u>	}	2 3	; ;		E		×	c	-	>	Н	×	H	æ	M	a
307	→	n	×	<	۷	E	2	2	•	4	:	,	,	,	,					l

FIGURE 9C

979 327	TITGAGAATAAAACAATAGTCTTTAATCACTCCTCAGGAGGGGACCCCAGAAATTGTAATG F E N K T I V F N H S S G G D P E I V M
1039 347	CACAGTTTTAATTGTGGAGAGAATTTTTCTACTGTAATTCAACACAACTGTTTAATAGT H S F N C G G E F F Y C N S T Q L F N S
99	ACTTGGAATAATAATAGGGTCAAATAACACTGAAGGAAATACTATCACACTCCCA T w n n t e g s n n t e g n t i t l p
59 87	TGCAGAATAAAACAATATAAACATGTGGCAGGAAGTAGGAAAAGCAATGTATGCCCCT C R I K Q I I N M W Q E V G K A M Y A P
19	CCCATCAGAGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRGQIRCSSNITG
79	ggtggtattaatgagatcgagatcttcagacctggaggaggagatatgagggac g g i n e n g t e i f r p g g g d m r d
0 L	AATTGGAGAAGTGAATTATAAATATAAAGTAGTAAAATTGAACCATTAGGAGTAGCA N W R S E L Y K Y K V V K I E P L G V A
99 87	CCCACCAAGGCAAAGAGAGTGGTGCAAAGAGAAAATGAGCGGCCCCCACCAAGGAGAAGAGTGGTGCTGCAAAGAAAATGAGCGGCCCCCCACCAAGAGAAAATGAGCGCCCCCCCACCAAGAGAGAAAATGAGCGCCCCCCACCAAGAGAAAATGAGCGCCCCCCCAAGAGAAAATGAGCGCCCCCACCAAAGAGAGAG

ш

G

ဟ

လ

လ

z

Z

ഗ

ഗ

Z

H

Z

H

4

Z

ၒ

H

Ω

ပ

×

S

397 133

22/42

FIGURE 10A

FIGURE 10C FIGURE 10A FIGURE 10B

LAI AV3-CD4"

ATGGATGCAATGAAGAGAGGCTCTGCTGTGTGCTG CTGCTGTGTGGAGCAGTCTTCGTTTCGCCCAGCCAGGAAATCCATGCCCGATTCAGAAGAGGGGCCAGAACA TCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCTGTGTACCCACAGACCC **AACCCACAGAAGTAGTATTGGTAAATGTGACAGAAATTTTAACATGTGGAAAAATGACATGGTAGAA**CAG atgcatgagatataatcagtttatgggatcaaagcctaaagccatgtgtaaaattaacccactctgtgtt <u> Agtitaaagtgcactgatitggggaatgctactaataccaatagtagtaataccaatagtagtagtagcggggaa</u> Signal cleavage NarI ပ O Σ **&** ы × × > ⋖ H U ഠ I Z ۵, × > ш Z H z O Ы S ۵, X a H Δ G > > M Z 3 ဟ > > H Þ H Δ -1 S > H × H ∢ • 4 Н ப × Δ 4 O ίωÌ Ω ഗ 109 253 85 325 37 37 109 181 61

SHEET (RULE

FIGURE 10B

860	ATC.	BATC	.AT	3	3	5			88	8 8		ູ້			B 47 B	٤	ָּ ֖֖֖֖֖֖֖֖֓֞	(ָ ֖֖֖֖֖֖֖֖֞	g.	6	((Ì	6
157	Σ .	Œ.	≥ E	[2] [3]	25	ن ا	(P)	63	, 3 8 ⊢1	X) ; Z	ຸ້ ບ	် လ	§ • 6 20	\$ \$ æ	} } }	ာ် ကို		ာ လ	₹	§ 2 2 2 8 2 8 2 8 8 8 8 8 8 8 8 8 8 8 8	၌ ဖ	\$ \$ \$) 2 2	M M M E K G E I K N C S F N I S I R G K V Q
583	AAS	SCAL!	}TA	TGC	AFF	'FF	PRF	\TW	25g	TTG	ATA	T	TAC	3	TAG	AT.	LATO LATO	+ 1888	CT.	ີ່ວູ	ည်	R. R. T.		T T (4 C
181	25	[2]	≫	< ✓	الايع	6 FZ7)	~ ~	 9	34	73	۵	₩	×	<u>O</u> .,	⊶	Ω	æ	Q		} ⊱	ဟ	}	}	<u>.</u>	KEYAFFYKLOIIPIONDITSYLLT
613	₽ ©	Frci	S. S	CAC	CZC	.MGI	ig B	FTM(2% 2%	₽	C C	GTC	25	95 S	TAI	ເດ	TLL	. 38 kg	ည္တ	arr(ວິ	AT.A	CAI	7. A.Y.	rtgr
208	S	ပ	æ	L	(v)	منین	7	<u>. </u>		ō	æ	ບ	Q .,	×	A	က	[cz	[62]	Ω	⊶	۵	⊶	æ	≻	S C N T S V I T Q A C P K V S F E P I P I H Y C
68 8 8	පු	ງລູ	38	T.	TL	38 38	GA	rrc.	T.A.A.	AAT	GTA	ata	AT.	JAG.	ເດີ	ຳໄດ້	uat c	- 28 28 28	ACA(3G&C	ຊື່	rcr	₩ SC	3	CTC
229	æ	<u>0</u>	æ5	G	<u>S</u>	C	~		ر د	25	ပ ပ	22.	æ	쏦	⊱	(C24)	22	ල	€ ⊸	ဖ	Q .,	ပ	ૄ ⊸	Z	APAGFAILKCNNKTFNGTGPCTNV
757	35		K	WCD WCD	J.L.	BRE	30% (2)%	*3G	. S.	TTA	Š	CAG	Tag	THI	C.	CTC	SAS C	;TGC	C T G	rtg/	A.T.	98	ag.	CLI	ISCA
253	ഗ	₩	∌		O	رسي	معه	-		₩	Œ	Q _u	>	A	ശ	⊱	0	چَ	~	~	Z	ဗ	ഗ	ـ ـ	S T V O C T M G I R P V V S T Q L'L L M G S L A
65 C	20	Sec.	8	SE T	ragi	ing.)&Ri	MT	J.	් ට්	ATT	\mathfrak{IC}_{β}	5	B.C.B	ST.	CI	JANA,	S C C	TW	ata.	3TA(B	CIG	Ž	SC &
277	(M)	Wa .	(M)	A	₽ ′	~	~~	6£	رم درم	æ	22	معا	£	Q	æ	æ	25	⊱	⊷	~ ⊸	>	Ō	~3	Z	E E E V V I R S A N F T D N A K T I I V O L N O
و 2	J.	rer.	8	200 T	T. M.	\P56	FR. P.C.	. B	3	C. See	GAC.	B.P.B	Gr.	E	TL	GT?	GAG	S.		EG	MT(ည္တ	&CT	111	AS S
301	ွ်တ	3	48€	≈	226		ירט	<u>~</u>	ແລ	≪(©	6	ပ	22	~	S	œ	æ	2 5	33	Z	æ	⊱	⊸ .	S V E I N C 7 G A G E C N I S R A K W N A T L K
67.0	و	8 L & C	ر پي	7. B.G.		AT,	Z A A C	386	Š	ZAT.	TT	%	AT.	ATA		S _S	TAR	TCI	LLI	છ ્ક	**	ည္သင္	TCA	છુ	9
325	jo			S	35		, 62s	~	€a]	0	مما	ဖ	Z	Z	25	⊱	~	₩	لمعا	×	ō	လ	S	S	QIASKLREQFGNNKTIIFKOSSGG

NotI coeccec

1447

FIGURE 10C

GACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGGAATTTTTCTACTGTAATTCAACACAACTGTTT **AATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACAC**TGAAGGAAGTGACACAATCACA CTCCCATGCAGAATAAAACAATTTATAAACATGGTGCAGGAAGTAGGAAAAGCAATGTATGCCCCTCCCATC **AGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAA**GAGATGGTGGTAATAACAACAAT **GGGTCCGAGATCTTCAGA**CCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAA GTAGTAAAAATTGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAAGAGAGTGCAGAGAGAAAATGA Z Ω Z ഗ ပ . တ a ပ Σ G Ы Z K ပ Ω ဟ 노 œ æ æ G Z H 3 æ H z > z × H ဟ M 0 M ~ H O œ ပ 0 × ပ E ပ (L) H ပ Σ H Ω Q, G z ഗ Z H ~ Ĺ z S G 3 H > G S Ĺ., G ဟ 0 I S O. × ပ æ H Z > Н æ Ĺ., M H æ H H a ت ပ Ĺ ¥ > G S S Ω ഗ 1045 349 1189 1333 1405 1117 373 445 469 397 1261 421

24/42

			•		
	•				
	,	•			
-					
				·	
,					
		•			
	٠,				
	,				

$\mathbf{\omega}$
T-
T-
ш
$\mathbf{\alpha}$
\supset
<u> </u>
正

AGCATAAGAGATGAGGAAAGAATATGCTCTTTTTATAAACTTGATGTAGTACCA S I R D E V Q K E Y A L F Y K L D V V P ATAGATAATAATAATAATAATACCGCTATAGGTGATGTGACCTCAGTCATTACACAG I D N N T S Y R L I S C D T S V I T Q GCCTGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT A C P K I S F E P I P I H Y C A P A G F GCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAAATGGTAAATGGC A I L K C N D K T F N G K G P C K N V S ACAGTACCAAAGAAAATGAAATTAGATCTGACAATTCAGAACTGCTAAATGGC AGTCTAGCAGAAGAAGAATTAGATCTGACAATTTCAGAACAAATGCTAAAATGGC AGTCTAGCAGAAGAAGAATTAGATTAG	ATAGGAGATGAGAATATGCTCTTTTTATAAACTTGATGTAGTACCA I R D E V Q K E Y A L F Y K L D V V P NGATAATAATAATACCAGCTATAGGTTGATAAGTTGTGACCTCAGTCATTACACAG D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATTAGGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S CT H G I R P V V S T Q L L L N G CTAGCAGAAGAAGAATTAGATTAGATCTGACAATTCACGAACAATGCTAAAACC L A E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAATTAATATTGTACAGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N AGTAGAACAAATGGAATTAAAATTAAAAAATTAAAAAAAA		GATAGC	GAG	(5)	ည္က	ATG :	GAG 1	AGA 1	5 5 (SA I	ATA	₹:	3	71.00 1.00 1.00 1.00 1.00 1.00 1.00 1.00	E I	TIC	¥.	[AT	CAC	CAC
ATAGAGATGAGATATGCTCTTTTTTATAAACTTGATGTAGTACCA I R D E V Q K E Y A L F Y K L D V V P GATAATAATAATAATAACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S GY Q C T H G I R P V V S T Q L L L N G CTAGCAGAGAAGAAGAGACTTGAAATGCTAAAAACC CTAGCAGAAGAAGAAGAATTAGATTAATTAATTGAAAATGCTAAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAATTAATAATTGTAACAGGACCATGTAAAAC I V Q L K E S V E I N C T G A G B C N	ATAGAGATGAGGTGCAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCAG I R D E V Q K E Y A L F Y K L D V V P GATAATAATAATAACCAGCTATAGGTTGATGACACCCCAGCTCATTACACAG D N N N T S Y R L I S C D T S V I T Q TGTCCAAAGATACCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S GYACAATGTACACATGGAATTAGACCAGTAGTAACACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L D N G CTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAACTACTGCTAAATGGC ATAGTACAACAGAAAAAAATGAAAATTAGAATTATAAAAAATTAAAAAAA		S	ы	ს	H	Σ	ы	æ	ပ	M	Н	×	Z	ပ	လ	[LL	Z	H .	H	H
I R D E V Q K E Y A L F Y K L D V V P GATAATAATAATACAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAAATGTCAATGCCAGTTCAATGGAAAAGGACCATGTAAAAAATGCTAAATGGC I L K C N D K T F N G K G P C K N V S GYACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAACTGCTAAATGGC V Q C T H G I R P V V S T Q L L N G CTAGCAGAAGAAGAGGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAATTATAATTGTACAGGTGCTGGAACATTGTAAC I V Q L K E S V E I N C T G A G B C N	I R D E V Q K E Y A L F Y K L D V V P PGATAATAATAATAATAATAATAATAATAATAATAATAATA		ATA	AGA	GAT	GAG	GIG	Ş	*	3 5	ITAT	SCI	CII	TII	TA	Z.	CT	3	ָ ב <u>ַ</u>	T.C.	7
GATAATAATAATACCAGCTATAGGTTGATAAGTTGTGACCTCCAGTCATTACCAGG D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAATGCCAGT I L K C N D K T F N G K G P C K N V S CTAGCAAGAAGAAGAAATTAGGCCAGTAGTACACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G CTAGCAAGAAGAAGAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E V V I R S D N F T N N A K T I V Q L K E S V E I N C T G A G B C N	GATAATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACTTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F I L K C N D K T F N G K G P C K N V S GYACAATGTACATGGAATAGGCCAGTAGTATCAACTCACTGTAAAAAGGC V Q C T H G I R P V V S T Q L L N G GTAGCAGAAGAAGAAGAATTAGATTAGATCTGACAATTCACGAACAATGCTAAAACC L A E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAGATTAATTGTACAGGTGCTGGACATTGTAAA I V Q L K E S V E I N C T G A G B C N S R A K W N D T L K Q I V I K L R E Q S R A K W N D T L K Q I V I K E B		H	a	۵	ы	>	0	×	ы	>	~	H	E.	>	×	13	Δ	>	>) a,
D N N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S GOVACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L D N G CTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATCTGTAGAATTAATTGTAACAGTGCTGCAAAAACC I V Q L K E S V E I N C T G A G B C N	D N N T S Y R L I S C D T S V I T Q TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F ATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S GYACAATGTACACATGGAATTAGGCCAGTAGTATCAACTGCTGCTAAATGGC V Q C T H G I R P V V S T Q L L D N G CTAGCAGAAGAAGAATTAGATTAGATCTGAAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAATTGTACAGGACAATGTAAA I V Q L K E S V E I N C T G A G B C N STAGAAAATGGAAATGACACTTTAAAAATTAAAAATTAAGAACAA S R A K W N D T L K Q I V I K L R E Q S R A K W N D T L K Q I V I K E B		GAT	AAT	AAT	MAT	N N	AGC	TAT	8 6	TTG	ATA	AGT	TGT	SAC	SS S	ZI:	GTC	AT	TAC	ACA
C P K I S F E P I P I H Y C A P A G F SATICTAAAGTAACCTTTGAGCCAATTCCCATACTTATTGTGCCCCGGCTGGTTTT SATICTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S NG!ACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCAGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G CTAGCAGAAGAAGAGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T LAAGTACAGCTGAAAGAATCTGTAGAATTAATTGTAACAGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N	C P K I S F E P I P I H Y C A P A G F SATICTAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTT C P K I S F E P I P I H Y C A P A G F SATICTAAAGTGTAATGGTAAGGCTTCAATGGAAAGGACCATGTAAAAATGCAGC I L K C N D K T F N G K G P C K N V S AGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L N G CTAGCAGAAGAAGAGGTAGTAATTAGATCTGAAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATTAAATTAATTGTACAGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N TAGTAGAGCAAAATGGAAATTAAAAATTAAAAATTAAAAAATTAAGAACAA S R A K W N D T L K Q I V I K L R E Q S R A K W N D T L K Q I V I K L R E Q		Ω	Z	Z	Z	H	ဟ	>	æ	H	H	လ	ပ	Ω	H	တ	>	H	€ →	a
C P K I S F E P I P I H Y C A P A G F SATICTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S GYACAATGTACACATGGAATTAGGCCAGTAGTATCAACTGCTGCTAAATGGC V Q C T H G I R P V V S T Q L L N G CTAGCAGAAGAAGAAGTTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T L A E E S V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N	CCPKISTER SFER TO THE TOTAL OF THE TOTAL OF THE TOTAL OF THE NOR TOTAL OF THE TOTAL OF TOTAL OF THE TOTAL OF TOTAL OF THE TOTAL OF TOTAL OF THE TOTAL OF T		TGT	ర్ట	AG	ATA	ည်	TTT	GAG	Ş	ATT	ည	ATA	CAT	TAI	TG	ပ္တ	ÿ	ည်	1 55	TIT
SCGATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAAGGACCATGTAAAAATGTCAGC A I L K C N D K T F N G K G P C K N V S CAGNACAATGTCAGCT CAGTACTACTCAATGTCAATGGC T N G T P V V S T Q L L L N G T V Q C T H G I R P V V S T Q L L L N G T V Q L L L N G T V Q L L L N G T V Q L L L N G T V Q L L L L N G T V Q L L L L N G T T V Q L L L L N G T V Q L L L L N G T V Q L L L L N G T V Q L L L L N G T V Q L L L L N G T V Q L L L L N G T V TAATAGTACAGAGAAAAAAAATTAATTGTAAATGTAAAAAGAAATTAATAGTACAAGAAAATTAATAGTACAAGAAATTAATAGTACAAGACAATTGTAACTAAAAAGAAAATTAATAGTACAAGAAAATTAATAGTACAAGAAATTGTAACTAAAAAAAA	GATTCTAAAGTGTAATGATAAGACGTTCAATGGAAAGGACCATGTAAAAATGTCAGC I L K C N D K T F N G K G P C K N V S AGTACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G CTAGCAGAAGAAGAGTAGTAATTAGATCTGACATTTCACGAACAATGCTAAAACC L A E E V V I R S D N F T N N A K T L A E S V E I N C T G A G H C N I V Q L K E S V E I N C T G A G H C N TAGTAGAGCAAAATGGAATGACACTTTAAAACAGATGTTATAAAATTAAAAATAAAAAAA I V Q L K E S V E I N C T G A G H C N TAGTAGAGCAAAATGGAATGACACTTTAAAACAGATGATATAAAATTAAAAATTAAAAAAAA		ပ	۵ı	×	H	S	(LL	ப	۵.	H	م	H	x	>	ပ	~	Δ,	K	G	Ŀ
I L K C N D K T F N G K G P C K N V S AG!ACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G TCTAGCAGAAGAAGAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T AATAGTACAGCTGAAAGAATTAATTAGTACAGGTGCTGGAACATTGTAAC I V Q L K E S V E I N C T G A G B C N	ACAGYACAATGAACTAGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC T V Q C T H G I R P V V S T Q L L L N G AGTCTAGCAGAAGAAGAAGTAATTAGATCTGACATTTCACGAACAATGCTAAAACC S L A E E E V V I R S D N F T N N A K T I I V Q L K E S V E I N C T G A G B C N ATTAGTAGAAAATGGAATGAATTAAATTGAACAGACAATTGTAAC I I V Q L K E S V E I N C T G A G B C N ATTAGTAGAAAATGGAATGAATGAAAATTAAAAATTAAGAAACAA ATTAGTAGAAAATGGAATGAATGAAAAAAAAAA	G	ATT	CTA	3	TGT	AT	GAT	MG	ACG	TTC	AAT	66	3	35	$\ddot{\mathcal{O}}$	vīgi	Z.	Š	TGT	S S
AGYACAATGTACACATGGAATTAGGCCAGTAGTATCAACTCCAACTGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G TCTAGCAGAAGAAGAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATCTGTAGAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N	AGYACAATGTACCATGGAATTAGGCCAGTAGTATCAACTCCTGCTAAATGGC V Q C T H G I R P V V S T Q L L L N G TCTAGCAGAAGAAGAAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T AATAGTACAGCTGAAAGAATCTGTAGAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N TAGTAGAGCAAATGGAATGACACTTTAAAACAGATATAAAAXTTAAGAGAACAA S R A K W N D T L K Q I V I K L R E Q		H	ı	×	ပ	Z	۵	×	H	Ĺ	Z	ပ	×	ပ	۵	ပ	×	Z	>	S
V Q C T H G I R P V V S T Q L L L N G TCTAGCAGAAGAAGAGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T NATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N	V Q C T H G I R P V V S T Q L L L N G TCTAGCAGAAGAGGGTAGTATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T AATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G H C N TAGTAGAGCAAATGGAATGACACTTTAAAACAGATAGTTATAAAATTAAGAACAA S R A K W N D T L K Q I V I K L R E Q	_	G7.7	3	TGT	Z	CAT	SGA GGA	ATT	AGG	₹ U	GTA	GTA	TCA	S	Ş.	CTC	SCI	SCT	\$	1 66
TCTAGCAGAAGAAGAGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T NATAGTACAGCTGAAAGAATTAATTAATTGTACAGGTGCTGGAACATTGTAAC I V O L K E S V E I N C T G A G B C N	TCTAGCAGAAGAAGGGTAGTAATTAGATCTGACAATTTCACGAACAATGCTAAAACC L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V O L K E S V E I N C T G A G B C N IAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAAAAATTAAGAAACAA S R A K W N D T L K O I V I K L R E O		>	0	ပ	H	×	U	H	∝ ′	۵	>	>	လ	H	0	H	.	-1	Z	G
LAEEEVVIRBABATTAATTGTACAGGTGCTGGACATTGTAAC NATAGTACAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAA	L A E E E V V I R S D N F T N N A K T ATAGTACAGCTGAAAGAATCTGTAGAATTAATTGTACAGGTGCTGGACATTGTAAC I V O L K E S V E I N C T G A G H C N TAGTAGAGCAAATGGAATGACACTTTAAAACAGATGATATAAAAATTAAGAGAACAA S R A K W N D T L K O I V I K L R E O	_	CTA	Ş	X 5	₩	SAG	GTA	GTA	ATT	AGA.	TCT	Sec	MI	TT	SKC SKC	Ž	₩.	ည်	¥	A C
NATAGTACAGCTGAAAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V Q L K E S V E I N C T G A G B C N	AATAGTACAGCTGAAATCTGTAGAAATTAATTGTACAGGTGCTGGACATTGTAAC I V O L K E S V E I N C T G A G B C N TAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAAAATTAAGAGAACAA S R A K W N D T L K O I V I K L R E O	•	7	4	ш	ш	ш	>	>	H	æ	S	Ω	Z	(e,	۲	Z	Z	«	×	H
I V O L K E S V E I N C 7 G A G B C N	I V O L K E S V E I N C T G A G H C N TAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAAAATTAAGAGAACAA S R A K W N D T L K Q I V I K L R E Q		ATA	GTA	9	CTC	₹ ¥	8	ICI	GTA	SA SA	ATT	AT	161	ğ	S	ទ្ធ	3	ğ	rtg	Ž
	FAGTAGAGCAAAATGGAATGACACTTTAAAACAGATAGTTATAAAATTAAGAGAACAA S R A K W N D T L K Q I V I K L R E Q	-	H	>	0	7	×	M	S	>	ធ	H	z	ບ	E 4	G	4	U	Ħ	ပ	Z
			လ	«	K	×	3	z	Ω	H	7	×	>	-	>	→	<	4	K.	a	>

FIGURE 11C

45		_		_		4	₹.
ATG	AGT S	6 00	CCT	GAT	GAC	AGC	
GTA V	Z	CH CH	ပ္ပ ∢	AGA R	A GG	AGT V	
ATT(I	F	T P	X	NO F	M	ပ္ပိ ဝ	
SAA E	TGI	H C	KTGJ	Ž 1	SAT	L	Ŋ
CAC	- 2 G	T T	2 ≤	. H. H.	S 5) 	3
SACC	T	ATA	X X	1 1	GAG G	GAAC E	<u>8</u>
ည်	S	ည် ပို့ပ	၌ ပွဲ ပ	ပ္ပံ့ပ	GAG G	AII	GAG -
ည်တ	ATT	AAG E	TAG V	ر ۱ ۲	CTG P	3 ×	X
S S	GTA	r T	E	I I	IGAC R	GTA V	A H
က် လ	ACT	Z Z	986	ATA	TC.	GTA V	GAG
E H	TCT	ATA	6 >	S S	HCH	¥ ×	30
N	TTT F	ညီ က	7 X	CAT	AGA E	TAT	750
F	AAT	6 61 6	A CA	GTT	F T	₹×	7 V
7. V	GA G G	E E	XI H	GAT	668 6	TAT	GAG R
I	GAG G	T	I	I	ATG	L	₽,6 %
A F	ិ ក	ATA	30	30	E E	ES a	AGA K
GAATAAAACAATAGTCTTTAATCACTCCTCAGGAGGGGACCCAGAAATTGTAATG	ATT	AATAATAATACTGAAGGTCAAATAACACTGAAGGAAATACTATCACACTCC N N N T E G S N N T E G N T I L	AATAAAACAAATTATAAACATGGTGCAGGAAGTAGGAAAAGCAATGTATGCCC I K Q I I N M V Q E V G K Å M Y A	ည် ၁	TATTAATGAGAATCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGAC I N E N G T E I F R P G G G D M R D	GGAGAAGTGAATTATAAATATAAAGTAGTAAAATTGAACCATTAGGAGTAGCA W R S E L Y K Y K V K I E P L G V A	CAAGGCAAAGAGAAGTGGTGCAAAGAGAAAAATGAGCGGCCGC
ATA	TTA	ATA	A I	68 6	TTA	AGA R	≯
AGA	GTT S	¥ GGA	2 8	TCA	GTA G	7GG ¥	S F
TTTGA F E	CACAGITITAAIIGIGGAGGAGAAIIIITCIACIGIAAIICAACACAACA	ACTIGGAATAATAATGAAGGGTCAAATAACACTGAAGGAAATACTATCACACTCCCA T w n n n t e g s n n t e g n t i t p	TGCAGAATAAAAATTATAAACATGGFGCAGGAAGTAGGAAAAGCAATGTATGCCCCT C R I K Q I I N M V Q E V G K A M Y A P	CCCATCAGAGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACAAGAGAT PIRGQIRCSSNITGC	GGTGGTATTAATGAGATCGAGATCTTCAGACCTGGAGGAGATATGAGGAC G G I N E N G T E I F R P G G G D M R D	AATTO	CCCACC
979 327	1039 347	1099 367	1159 387	1219	1279	1339	399 487

FIGURE 12A

LAI CD4

FIGURE 12A FIGURE 12B

FIGURE 12C

ATGGATGCAATGAAGAGAGGGCTCTGCTGT tPA signal sequence

Nari ပ ¥ **∑**

G × Ř ⋖ I Н ш O ⋖ O ပ ,1 13 37

H H Ш × 3 > ۵, > G > > H > 3 Ы 109 37

TGTGCATCAGATGCTAAAGCATATGATACAGAGGTACATAATGTTTGGGCCACACATGCTGTGTACCCACA ىم > ပ K = H 4 3 > z = > M H Ω Æ Signal cleavage × K ഗ K 181 61 gaccccaacccacaagaagtagtattggtaatgtgacagaaaattttaa¢atgtggaaaaatgacatggta E Z بما Z 回 H > Z > H > > W z 253 85 GAACAGATGCATGAGGATATAATCAGTTTATGGGATCAAAGCCTAAAGCCA'TGTGTAAAATTAACCCCACTC ပ × H S 0 0 S 325 109

tgtgttagtttaaagtgcactgatttggggaatgctactaataccaatagtagtaataccaatagtagtagt S ဟ S Z z H 4 Z G 0 H ပ × . ഗ > ပ 397 133

29/42 Ctagagaagaggtagtaattagattacccaatttcacagecaatgctaaaaccataatagtacaktt aaccaatctgtagaaattaattgtacaagacccaacaatacaagaaaaaagtatccgtatccagaggga CCAGGGAGAGATTGTTACAATAGGAAAATAGGAAATATGAGACAAGCACATTGTAACATTAGTAGAGCA GTGCAGAAAGAATATGCATTTTTATAAACTTGATATAATACAATAGATAATGATACTACCAGCTATACG tattgtgcccccccctgcttttccaatctaaatgtaaaraacaccttcaatgaacagccatgtaca antercaccacacaticanteracatacatiscantiagecageastactaracteaaticaatiscast **GGGGAAATGATGATGGAGAAAGGAGAGATAAAAACTGCTCTTTCAATATCAGCACAAGCATAAGAGGTAAG** ttgacaagttgtaacetcagtcattacagecctgtccaaaggtatcctttgagccaattcccatacat ഗ ပ Œ ဟ Z چ, ပ ဗ Ω (J) 0 × 天 - . 心 Z Z (co Æ ഗ Œ Z Œ O ഗ Þ Z **~** Q D œ 幺 26 Œ (CL) Q, Þ ⊱ Z Z \mathbf{x} ഗ **~** Q., Q. Z ပ Z مما Z ပ **~** ≪ ပ œ Z Z ල Z Æ ټ. 0 丝 Q. يخ ල ہہ ഗ Œ 幺 ڪ يج ₩ Œ Z G æ (A) Œ چ ပ ⋈ G ⋗ محا ပ Z Þ (CO) (Co. يم G Ö ð Ð Œ Ŋ W <u>[6</u>2] ശ Z Œ ≫ ⊱ U ഹ്ര Œ പ്ര ശ ശ 丒 <u>(M</u> G Þ 0 G Z ۵ 829 973 685 229 757 253 277 903 301 325 469 613 208 157 543 181

FIGURE 12C

30/42

TTTAAGCAATCCTCAGGAGGGACCCAGAAATTGTAACGCACAGTTTTAATTGTGGAGGGAATTTTTTACTAC GATGGTGGTAATAACAACAATGGGTCCGAGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGA **GCAATGTATGCCCCTCCCATCAGCGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAA**CAAGA TGTAATTCAACACAACTGTTTAATAGTACTTGGTTTAATAGTACTTGGAGTACTGAAGGGTCAAATAACACT **AAATGGAATGCCACTTTAAAACA**GATAGCTAGCAAATTAAGAGAACAATTTGGAAATAATAAAACAATAAT GAAGGAAGTGACACAATCACACTCCCATGCAGAATAAAACAATTTATAAACATGGTGAGGAAGTAGGA:AA G .-1 æ G M ပ Σ Σ ပ ပ H z z S H ပ z G တ ப S H L G 0 ഗ æ I S × _ Z ပ H æ H ¥ > æ æ H ပ 0 Ы H Ш ۵. Δ, S ပ Ś ... 0 တ ပ Ω Z H ပ × z بعا H H ပ . م Z H 0 H S Z م **_** K 4 z တ ڻ 0 ഗ ပ **U** Z M 1117 1045 349 373 1189 1333 445 1405 469 397 1261 421 SUBSTITUTE SHEET (RULLE 26)

ပ H Ш × > > × × × H 1477 493

· NotI

GTGGTGCAGAGAAAAATGAGCGGCCGC 1549

L æ 0 > 517

FIGURE 13A

FIGURE 13A

		GRAGAGA 1 K R	AGGAAATC
FIGURE 13B FIGURE 13C FIGURE 13D		ATGGATGCAATGAAGAGA M D A M K R	GGGCTCTGCTGTGTGCTGTGTGTGAGCAGTCTTCGTTTCGCCCAGCCAG
	CD4.		GGCTCTGCTGT

JR-FL

CATGCCCGATTCAGAAGAGGCGCAGAGTAGAAAAGTTGTGGGTCACAGTCTATTATGGG CAAGAAGTAGTATTGGAAAATGTAACAGAACATTTTAACATGTGGAAAAATAACATGGTA GILACAGAGGTACATAATGTTTGGGCCACACATGCCTGTGTACCCACAGACCCAAACCCA GTACCTGTGTGGAAAGAAGCAACCACCACTCTATTTTGTGCATCAGATGCTAAAGCATAT ۲ ۵ ဟ > ۰ × H ဟ S > Signal cleavage > U > X Z ⋖ 0 > 山 ပ r r Nari > H H ы Z Þ × H U ບ æ H ⋖ O 2 139 27 199 259 87 47 67

13

FIGURE 13E

						•
₹ X	×	ACCCCACTCTGTGTTAAATTGCAAGGATGTGAATGCTACTAATACCACTAAT T P L C V T L N C K D V N A T N T T N	agcgagggaacgagagagagaaaaaaaaagtgctctttcaatatcaccaca S E G T M E R G E I K N C S F N I T T	ATAAGAGATGAGGAGAAAGAATATGCTCTTTTTTATAAACTTGATGTAGTACCA I R D E V Q K E Y A L F Y K L D V V P	ATAGATAATAATACCAGCTATAGGTTGATAAGTTGTGACACCTCAGTCATTACACAG	TGTCCAAAGATATCCTTTGAGCCAATTCCCATACATTATTGTGCCCCGGCTGGTTTTCC PKISFE FIPIL TCCATACATTATTGTGCCCCGGCTGGTTTTT
GTA	>	ACT	ACC F	GTA V	Ž t	မှိနှင့် ၁
IGI	ပ	ACC T	ATC	GTA V	ATT	SCT A
S	Ω,	Z	Z Z	GAT	51.C >	ည္တည္
ტ ჯ	×	ACT.		C. T.	ည်လ	ည်း
7 7 7	H	3CT	S	¥ ×) 1	ត្តិ
ပ္တ	တ	ZZ	ည်ပ	rati Y	75 25 25 25 25 25 25 25 25 25 25 25 25 25	KAT X
3	ø	stg. V	S Z	rtti F	ည် ၁	H
34	A	D D	₹	L	S	ATA(I
ğ	3	§ ×	I I	3CTC	I	D N
TI	H	ပ္ပိုင	ξ E	rato X	រី ភូមិ	I
S	S	5 z	9 0 0	EX	1	3 ~
בְּל	H	ZI 1	NGA R	Ž×	TAT	SAG E
77	H	St.	3¥6.	30	ည့် ဖ	rtt(F
SAT	Ω	V	A X	37G *	ğ	ည်လ
S S	M	ည်ပ	્ર વ	3366	Z Z	HTA
S S	0	SH I	က္လွတ္	SAT	ZZ	बुँ×
S S S	X	N P	E S	कु स	Z Z	Ŋ.
P A	a	D F	S	MT.	D T	ည်ပ
3	E O M O E D I I S L W D O S L K P C V K	TIN	GNT	AGC. s	ATAC	& €
319	107	379 127	439	499	559 187	619 207

FIGURE 13C

လ ဗွင္မ	၁	∌cc ₹	Stacaectgaargetstacaattaattetacaacaacaataca v o l k e s v e i n c t r f n n n t	දී ය	ල ව
G 7C &	AAT R	A H	LAAT R	LATA	E
MAT N	CI L	ည် ﴿	SE S	aat I	rrr
raa K	FCT L	E R	SALAS SE	S	ر ال ال
ATG C	15 13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	ال ال ال	3 3 3	E C
ည္က	දූ ර	CAC T	AAG M	TAC	E
99 S	AAC F	TTI I F	Tac T	atac (I	LATO K
A A C	rato y s	25 C	arro K	TTT! F	Carl B
A T C	T & C	CT S	TTW I	Car B	ର୍ଜ୍ଞ ଅ
TCA F	282 8	Gat R	E S	S GAG R	SCTA
ACGT T	3 8 8 9	ATT I	37.20 V	මී ප	ATT <i>b</i> I
379 X	ATT/ I	GTW V	3CI(S	ည္တို 🏊	28 E
GAT	තු ව	Gra	E E	ම් ව	rrgr C
laat R	CAT R	rcac E	100 F	FAT? I	ACAJ ≊
FE C	RAC.	AGAN E	ICI K	ACA BE	300 es
ARAN K	ATG. C	MGA E	ACA O	TYL	ACA O
FCT/	SCE O	SSE CE	A A	JAJAG S	. A. B. C.
gcgnttctaragtgtaatgataacgtcaaaaggaccatgtaaaaatgtcagc a 1 l k c n d k t f n g k g p c k n v s	acastrcaatstacatrgcaattaggccastastrcaactgctstaatggc t v o c t k c 1 R p v v s t o l l l n c	ngtctagcagaagagggtagtagttgacaatttcacgaacaatgctaaagc s l a e e e v v i r s d n f t n n a k t	ataatagi III	rglalargetatrocaccaccaccatititatactrcaccacarataatagca r k s i k i c p c r a f y t t c e i i c	Gatatlagacarcattgtaacattagtagacaaragacactttaaaacag d i r q a h c h i s r a k w n d t l k q
679 (227	739 1	799	859 287	919	979

507

34/42

FIGURE 13D

1039	ATAGITATAAAATTAAGAGAACAATTIGAGAATAAAACAATAGICTTTAATCACTCCTCA	TAT.	₹.	¥,	Į.	KGA.	X	\$	TII	SAG	E :	\$:	ర్జ్ :	ATA	GTC:	TIT	AT.	: ن	ii ĭ	SH C	-
347	-	→		ب	,a	oc.	ia)	9	معا	ia)	z	×	H	H	>	ļae	Z	I	S	S	
1099	GGAGC	3GGA	ပ္ပံ	343	3	VTT(STA	ATG	CAC	AGT	TTI	Z	TGT	5 5	668	. 💆	TTI	TTC	TA	H	-
367	G G D P E I V M H S F N C G G E F F Y C	ري ص	_	ο.	ш	н	>	Σ	=	တ	Ĺ,	z	ပ	O	Ŋ		[a,	Ĺ	>	U	
1159	AATTCAACACACTGTTTAATAGTACTTGGAATAATAATACTGAAGGGTCAAATAACACT	CAAC	Ş	S	īĞ	TI	MAT	AGT	ACT	TGG	MAT	MI	AAT	ACT	35	- 13	Ş	3	Ž	5	H
387	z	E	_	C C	-1	[in	z	S	€ →.	3	Z	Z	z	H	Œ	<u>ල</u>	S	Z	Z	H	
1219	GAAGGAAATACTATCACACTCCCATGCAGAATAAAACAAATTATAAACATGGTGCAGGAA	SAMA	TAC	TA	J.	Š		CA	TGC,	PG	ATA	*	3	ATI	ATA	AAC	ATG	95	Ž	3GA	Æ
407	ы	G.NTITTPCRIKOIINMVOE	_	£	H	H	H	Q,	ပ	œ	н	×	0	H	H	2	Σ	>	Ø	Ы	
1270	GTAG	SAAA	ğ	3	IG	[AT	ပိုင်	Č	ပ္ပ	ATC	AGA	4 55	3	ATI	AGA	TGI	TC	Z C	3	IAT	F
427	V G K A M Y A P P I R G Q I R C S S N I	Σ Α			E	>-	×	ىم	Δ.	н	ac	ပ	0	H	œ	ပ	ဟ	တ	Z	H	
		נינים	ز	TAT	Ž		AGA(SAT(GGI	GGT	ATT	MAT	GAG	MAI	8	NCC.	SGAC	SAT	TT	CAG	4
447	TGLLLTRDGGINENGTEIFR	1) (S	} _	د	1	H	æ	۵	છ .	ပ	~	z	ы	Z	ပ	Ħ	M	H	(e)	oc.	
6		£7.	Z A	5 45	AT	VTG.	QQ.	3 A C.	MAT	TGG	AGA	AGT	3	TTA	ITAI	AN	LTAT	3	AGT.	AGT	«
1399	PGGGDMRDNWRSELYKKVV				۵	Σ	«	Ω	z	*	œ	လ	ы	H	⊭	×	×	×	>	>	
	A A A THER A CONTACTACTACTACCACCAAGGCAAAGAGAGAGAGTGGTGCAAAGAAGAA	£	Č	4	T.	AU.	TA)	Š	C)	Ç	Ŭ Ž	5 5	M G	AGA	AGA	GTC	SGTG	Š	25	AGA	æ
1459	AAAA x	5 K	ر د	4	֡֝֝֓֞֝֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֓֡֓֓֓֡	် ဗ	; >	4	۵,	H	×	~	×	æ	~	>	>	ø	æ	ы	
0	-	, -	NotI	. –)	•															•
1519	AAATGAGCGGCCGC	MGC	ğ	X	Ų																

FIGURE 14A

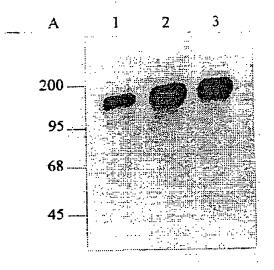


FIGURE 14B

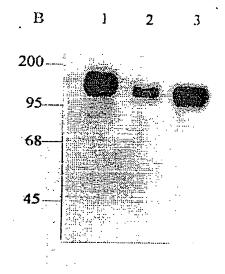
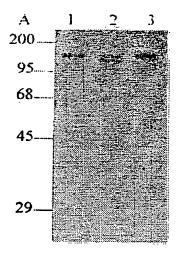


FIGURE 15A



PCT/US94/03282

38/42

FIGURE 15B

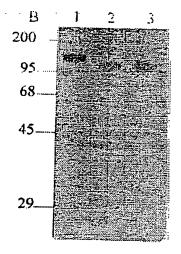
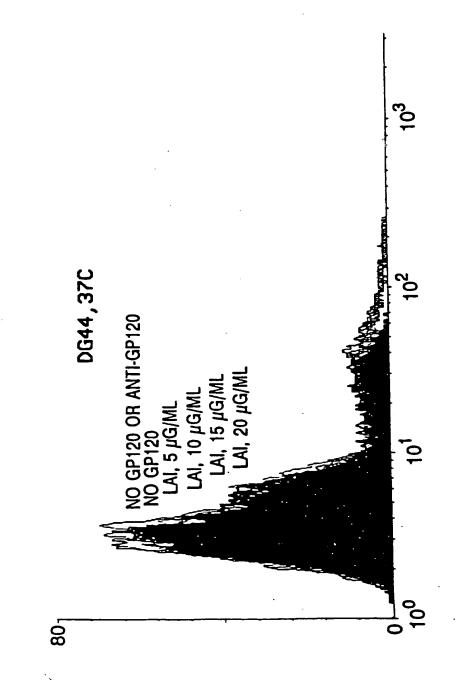
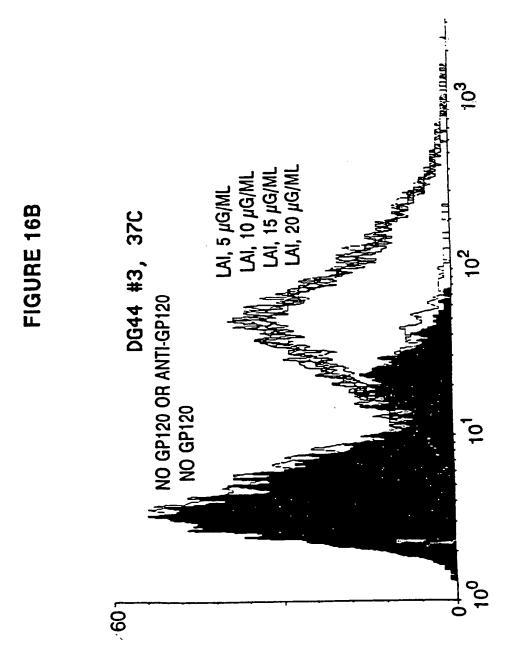


FIGURE 16A

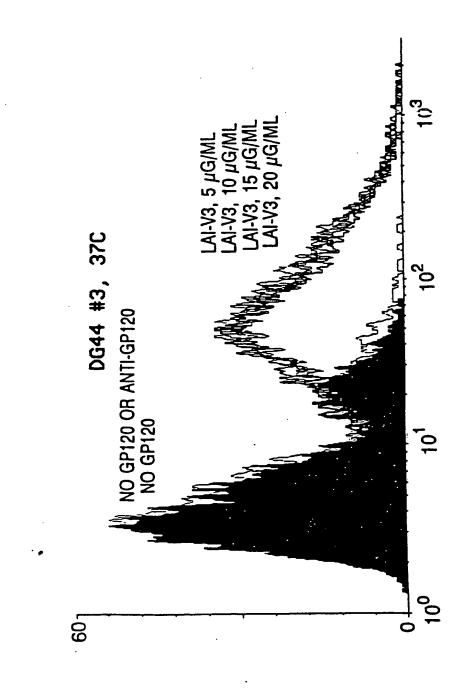


SUBSTITUTE SHEET (RULE 26)

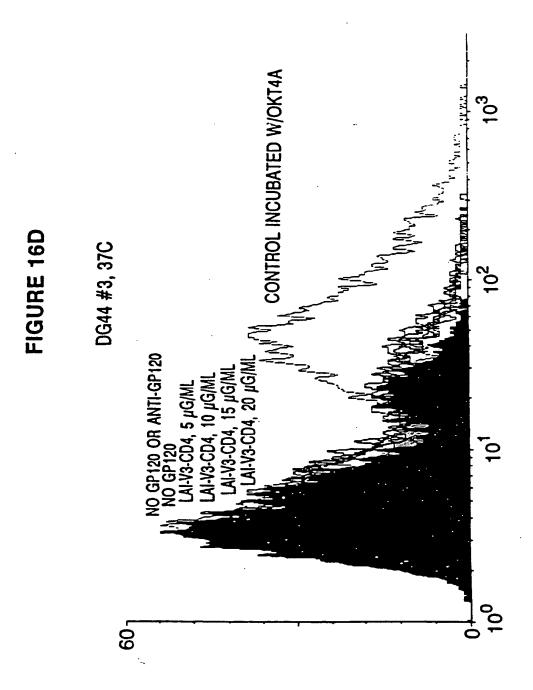


SUBSTITUTE SHEET (RIII F 26)

FIGURE 16C



SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

C:	A CCURRO A STANLA ME CUR DECOR DE A COMPANI						
IPC(5)	ASSIFICATION OF SUBJECT MATTER :Pleace See Enum Sheet.						
	: 424/88, 89; 536/27; 530/395						
	to International Patent Classification (IPC) or to i	both national classification and IPC					
B. FII	elds searched						
Meaninem	documentation pearched (classification system follo	owed by classification symbols)					
i	424/88, 89; 536/27; 530/395	- woo o, c	;				
0.0.	42.00, 07, 00021, 00021						
Document	ntion searched other than minimum documentation to	the extent that such documents are included	in the fields seembad				
			and the section of th				
Electronic	data base consulted during the international search	(name of data base and, where practicable,	search terms uned)				
APS, D	ialog, search terms: HIV-1, mutation, V3 loo	p, C4 region, envelope alycoprotein, w	occines puclaia anid				
			seemes, ribcleic acig				
C. DO	CUMENTS CONSIDERED TO BE RELEVANT	<u> </u>					
Colegoryo	Citation of document, with indication, where	appropriate of the relevant passages	Relevant to claim No.				
			Reteault to comm No.				
Y	Science, Volume 252, issued	17 May 1991, S. Wain-	6.7				
	Hobson, et al, "LAV Revisited:	Origins of the Early HIV-1	•				
	Isolates from Institut Pasteur",	pages 961-965, see entire					
	article.						
,	11C A F 000 440 4000						
Y	US, A, 5,030,449 (BERZOFSKY	ET AL) 09 July 1991, cols.	1-27				
}	4-7 .						
Y	NAIO A DAIAREAD (CDCCCO)						
'	WO, A, 91/15512 (GREGORY entire patent.	EI AL) 17 October 1991,	1-27				
1	entire patent.	i					
y	WO, A, 91/11461 (PASEK ET AL) 08 August 1991, see 1-27						
•	entire patent.	ALI US AUGUST 1991, See	1-27				
	omi potem.						
1							
1		•					
1							
1							
7	and a second						
	r documents are listed in the continuation of Box	C. See patent family annex.					
•	ind entergorism of criteri documents:	The document published after the inter- date and not in conflict with the applica-	national filing date or priority				
deru to be	ment defining the general costs of the art which is not considered of particular relevance	principle or theory underlying the inven	note that called to traderimed the				
eorlin	er document published on or after the international filing date	"X" document of particular relevance; the completed novel or cannot be considered	chimed invention connect be				
deac	throw doubte on priority chamte) or which is to combine the publication date of contact citation or other	when the document is taken olone	an to macuae on macoriae orch				
chacc	n record (co cpacified)	"Y" document of particular relevance; the	chimed invention cannot be				
), q eo n	ment referring to an oral disclarate, use, exhibition or other	combined to involve an inventive of	decrements, such combination				
	near published prior to the international filing date but buer than	being obvious to a person skilled in the					
Ext pe	tortly date classed	, , , , , , , , , , , , , , , , , , , ,					
are of the ac	tual completion of the international search	Date of mailing of the international conf	ch report				
09 JUNE 19	94	1 9917 2 4 1337					
	The sales of the s						
Commissiones	ling address of the ISA/US of Patents and Trademarks	Authorized officer	ma la				
Box PCT Wochington, I		LYNETTE F. SMITH	ga fr				
			- /				
esimile No.	(703) 305-3230	Telephone No. (703) 308-0196	•				

INTERNATIONAL SEARCH REPORT

International application No. PCT/US94/03282

A. CLASSIFICATION OF SUBJECT MATTER:		
IPC (5):		
	CO7H 15/12	13/00

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

IMAGES ARE BEST AVAILABLE COPY.

OTHER:

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.